

博士論文

保全緑地における生態系調査への深層学習法の応用研究

2023年3月

北九州市立大学大学院

国際環境工学研究科 環境システム専攻

中山 紘喜

目次

第1章 序論 1

1.1 研究の背景 1

1.2 研究の目的と構成 3

1.2.1 研究の目的 3

1.2.2 本論文の構成 3

参考文献 4

第2章 既往研究 5

2.1 深層学習法を用いた植物識別に関する既往研究 5

2.2 深層学習法を用いた鳥類検出に関する既往研究 5

参考文献 6

第3章 研究手法 9

3.1 畳み込みニューラルネットワーク (CNN) の概要と神経回路網モデルの性能 9

3.2 Keras による深層学習法の実装 10

3.3 Explainable AI による着目領域の可視化 11

3.4 MeanShift 法を用いたスーパーピクセル分割 12

3.5 YOLO 法による物体検出 12

参考文献 14

第4章 深層学習法による植物画像識別の植物調査への応用 15

4.1 はじめに 15

4.2 植物種画像データベースと学習モデルの構築手法 15

4.2.1 植物調査地域および調査の概要 15

4.2.2 深層学習方法 17

4.2.3 XAI 解析 17

4.3 調査・学習結果 17

4.3.1 2019 年 2 調査地点の深層学習結果 17

4.3.2 2019 年学習結果における確度の低い植物の要因解析 19

(1) クズ 19

(2) チガヤ 20

(3) タチスズメノヒエ 21

4.3.3 2020 年春季調査を追加した 2020 春モデルの深層学習結果 22

4.3.4 2020 春モデルにおけるヨモギの確度低下の要因解析とその改善 24

4.3.5 2020 年秋 3 調査地点の深層学習結果 27

4.4 考察 29

4.5 本章のまとめ 29

参考文献 30

第5章 深層学習法による植生分布の解析 31

5.1 はじめに 31

5.2 植生分布の解析手法	31
5.2.1 調査地区	31
5.2.2 深層学習方法	32
5.2.3 学習画像と試験画像の作成	32
5.3 調査, 学習結果および分布図の作成	33
5.3.1 木本種用画像データベースと試験画像の作成	33
5.3.2 学習と試験における確度	35
5.3.3 画像データベースの再構築と確度検証	37
5.3.4 樹木分布図の作成	39
5.3.5 俯瞰画像からの草本種識別結果	41
5.4 考察	42
5.5 本章のまとめ	43
参考文献	44
第6章 深層学習法による鳥類のリアルタイム動画解析	45
6.1 はじめに	45
6.2 鳥類自動検出, 追跡システム構築手法	46
6.2.1 撮影場所, 撮影機材	46
6.2.2 深層学習法の入力パラメータ, 神経回路網モデルの最適化	46
6.2.3 検出物追跡システム, 録画システムの開発	47
6.2.4 開発したシステムの常時運用	47
6.2.5 学習モデルの改善	47
6.3 システムの開発および常時運用	48
6.3.1 YOLO 法による鳥類自動検出	48
(1) パラメータの最適化	48
(2) 神経回路網モデルの最適化	48
6.3.2 追跡システム, 録画システムの開発	49
(1) 固定検出物の除去	49
(2) 検出物追跡システムの構築	49
(3) 検出時録画システムの構築	50
6.3.3 システムの常時運用	50
6.3.4 学習モデルの改善	53
6.4 議論	54
6.5 本章のまとめ	54
参考文献	55
第7章 議論	56
第8章 結論	58
8.1 各章のまとめ	58

8.2 本論文の結論	58
謝辞	60

1 章 序論

1.1 研究の背景

我が国では里山や都市部における河川，保全緑地等の人手によって管理されている自然が数多くあり，これらは生物多様性の保護だけでなく環境教育や住民の憩いの場としても重要である．北九州市では，撥川や板櫃川が従来の治水優先の河川改修から植生を伴う河川改修へ方針を転換した¹⁾．これにより川幅が広がり，複雑な流れが生まれ，転換前と比べ多くの環境が育まれるようになった（図 1-1，図 1-2）．板櫃川ではその後水辺の国勢調査や水辺の楽校が実施され²⁾，モニタリングや環境学習の場となっている．ひびきの北公園では，周辺の開発時に発見されたキンランやトラノオスズカケといった重要種を保全，繁殖させるための保全地区が整備された．またこの地区では絶滅危惧 2 類に分類されている，カスミサンショウウオの調査保全活動もなされている．響灘埋立地は重要種の保全地区と開発地区が隣接しており，当地区の環境保全及び生態系調査は都市部における自然環境管理手法において重要な意味を持つ．当地区で保全されている種の中には絶滅危惧種に分類されるチュウヒ（タカ科の鳥類，国・絶滅危惧 I B 類，国内希少野生動植物）やベッコウトンボ（昆虫類，国・絶滅危惧 I A 類，国内希少野生動植物）³⁾がいる．



引用) 北九州市：板櫃川（都市基盤整備事業）

図 1-1 板櫃川河川改修による河川環境の変化



引用) 北九州市：撥川（都市基盤河川改修事業）

図 1-2 撥川河川改修による河川環境の変化

このように響灘ビオトープをはじめ、都市部の自然環境や里山には重要種を含めた様々な生態系が形成されているが、多くの地区でその維持管理が課題となっている。中でも生態系のベースとなる植物種の種数や分布の把握、行動圏が広く生態系の頂点に位置する大型鳥類種の生活行動の把握は重要である。しかし、重要種保全の必要性が十分に認識されていないことに加え、維持管理のための調査では専門家や経験者が必須となり、その頻度や地域が制限されてしまう。このような要因が重なり、重要種の生息が確認されていても十分な生態系調査がなされていない地区も多い。北九州市では前述した響灘ビオトープ、ひびきの北公園、撥川において、生態系調査は不定期で行われており、定期的な調査は行われていない。これは全国の都市緑地に共通の課題であり、保全上重要なモニタリングの困難さを軽減する効率的な手法が必要とされている。

このような課題に対して、近年発展が著しい深層学習法や小型無人航空機 (UAV) といった IT の活用が有効と考えられる。深層学習法の応用は、その有用性が確認されて以来様々な分野に応用されている。深層学習法は人工知能の一分野である機械学習に属する手法で、機械学習において課題であった特徴量の抽出を、大量の入力データを用いて自動で行うことができる。生態系調査では主に視覚情報からデータを得るが、深層学習法の映像分野での活用は、画像と正解ラベルを紐づけた教師あり学習が盛んであり、何の画像であるか分類する物体分類、全体として何の画像であるかを認識するシーン認識、画像のどこに何があるか検出する物体検出、画像全体を物体で区分けする領域検出 (図 1-3) が挙げられる。中でも物体分類は 2010 年から 2017 年まで開催された画像認識コンテストの ILSVRC (ImageNet Large Scale Visual Recognition Challenge) ⁴⁾において、2012 年の AlexNet⁵⁾による 10%を超すエラー率の改善により深層学習法が注目されるきっかけとなった。このコンテストは ImageNet⁶⁾のデータセットを用いた 1000 種の分類を対象に、予測確率 5 位 (top5) までのエラー率を競っている。これは分類時に算出される各分類クラスの確からしさのうち、上位 5 位までに正解の分類クラスが含まれていない場合をエラーと見なした際の、全分類クラスの平均エラー率を表している。2015 年には ResNet (Residual Neural Network) ⁷⁾が 96.4%の top5 正答率を達成した。これは同じ条件での人間の平均 top5 正答率 94.9%を超えている。図 1-4 に ILSVRC の優勝モデルの top5 正答率の遷移を示す。



引用) Stanford background dataset

図 1-3 領域検出の一例

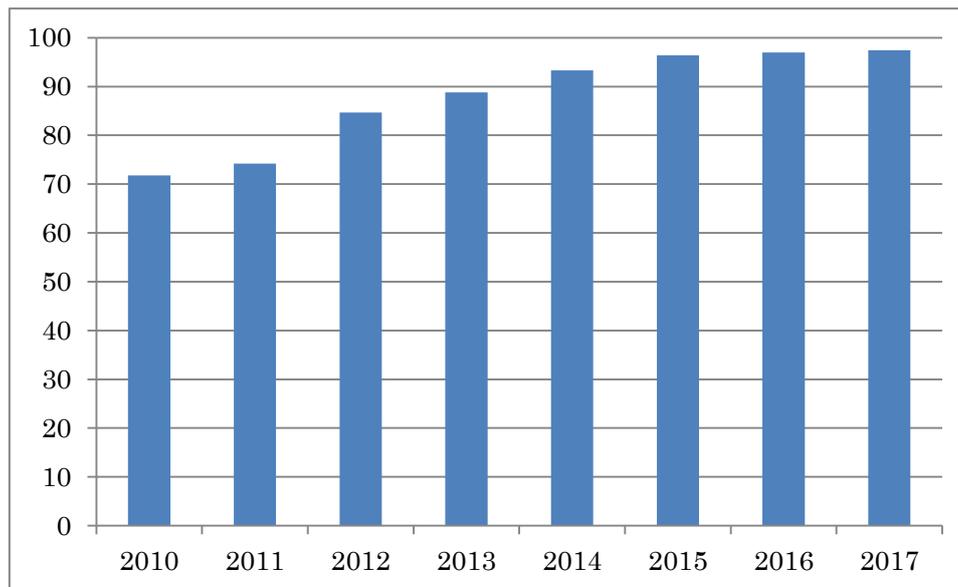


図 1-4 ILSVRC 優勝モデル top5 正答率の遷移

深層学習法による画像解析は自動車，医療，農業などの分野で応用研究が盛んであるが，写真画像からの種の識別，ビデオ映像からの種の検出，領域検出による植生図作成など，生態系調査への応用も期待できる．また小型 UAV はその扱いやすさと航続距離の長さから，農薬散布や被災状況の確認など幅広い分野で活用されている．生態系調査においては，森林上空や湿地帯など人の踏査が難しい地区での映像データの収集に有用である．以上より生態系調査において IT の応用，特に深層学習法の応用について研究することは大きな意義がある．

1.2 研究の目的と構成

1.2.1 研究の目的

北九州市若松区に位置する響灘ビオトープを中心に，保全緑地の生物種に対し深層学習法による教師あり学習を用いた生物調査手法を構築，運用しその適用性を評価することが本研究の目的である．具体的には植物と鳥類について，それぞれ深層学習法を用いたシステムを開発する．植物では複数の調査地点ごとにオリジナルの学習画像データベースを作成し，深層学習法による物体分類を用いて高精度植物識別システムを構築する．また，構築したシステムを用いてドローンで撮影した俯瞰画像から植物の分布を求める手法を構築する．鳥類では響灘ビオトープに設置された4Kネットワークカメラを活用し，深層学習法による物体検出を用いて，リアルタイムでの鳥類の自動検出および検出鳥類の自動追跡システムを開発する．また，開発したシステムの常時運用を行い，鳥類の生態解析へむけた検証とシステムの課題を明らかにする．

1.2.2 本論文の構成

本論文は，上記の目的に即した研究成果をまとめたものであり，全8章で構成される．第1章では，研究の背景について論じ，本研究の目的と構成を述べた．第2章では，深層学習法を用いた植物識別に関する既往研究，深層学習法を用いた鳥類検出に関する既往研究をレビューしたうえで，本研究分野の

現状を明らかにした。そのうえで既往研究の特徴や不足点等をまとめ、本研究の位置づけを明確にした。第3章では、本研究で引用した既存手法について、その概要を説明した。第4章では、草本類の調査により作成したオリジナル学習画像データベースと深層学習法Xceptionを用いて高精度な植物識別システムを構築した。その上で識別制度の低い種について、その要因を解明し制度改善を試みた。第5章では、木本類について第4章で最適化した学習パラメータを流用し、高精度な樹木識別システムを構築した。構築したシステムを用いて、ドローンで撮影した広域の森林空撮画像にたいし樹木の分布解析を行った。またドローンで撮影した草本類の俯瞰画像に対し第4章で構築したシステムによる植物識別と領域分割を用いて植生分布解析を行い、その適用性について検証した。第6章では、響灘ビオトープで撮影された動画を対象に、深層学習法YOLOを用いた鳥類の自動検出、自動追跡システムを構築した。構築したシステムは響灘ビオトープで常時運用を行い鳥類の生活行動解析について検証し、得られた成果と課題を明確にした。また得られた課題の解決策について検討した。第7章では、本研究で得られた知見を基にその将来性や展望について考察と提言を行った。第8章では、本研究で得られた知見を総括し、結論を述べた。

参考文献

- 1) 西野友子(2019) : 北九州市内河川公園の植物相とその特徴, わたしたちの自然史特別号, 初版, pp.31-62
- 2) 北九州市建設局河川部水環境課 (2016) : 平成27年度板櫃川(高見地区)水辺の楽校生態系調査ほか業務委託報告書
- 3) 環境省 (2020) : 環境省レッドリスト2020, <https://www.env.go.jp/press/107905.html>, 参照2023-01-06
- 4) LSVRC : <http://www.image-net.org/challenges/LSVRC/>, 参照2022-12-09
- 5) Alex, K., Ilya, S., and Geoffrey E. H. (2012) : Image-Net Classification with Deep Convolutional Neural Networks, In NIPS
- 6) ImageNet : <http://image-net.org/>, 参照2022-12-09
- 7) Kaiming, H., Xiangyu, Z., Shaoqing, R., and Jian, S. (2015) : Deep Residual Learning for Image Recognition, <https://arxiv.org/abs/1512.03385>, 参照2023-02-03

第2章 既往研究

2.1 深層学習法を用いた植物識別に関する既往研究

植物分野における深層学習法の応用研究としては、葉病理診断¹⁾、作物の機械的損傷の検出²⁾、種子画像による半数体の選別³⁾、花の図鑑アプリケーションの開発⁴⁾等がなされており、農業関連や図鑑利用での応用が多くなされている。植物識別では、Seeland 等が西欧の植物 1000 種の画像データベース PlantCLEF2016⁵⁾を用い、種で 82.2%、科で 88.4%の精度を得ている⁶⁾。深層学習法による画像識別を植物調査に適用できれば、専門家への負担も減り、定期的かつ包括的なモニタリング実施のハードルも下がると考えられるが、高精度が得られている ILSVRC の 1000 種の識別項目の中に植物は花を主として 10 数種しか含まれておらず、都市緑地での植物識別への利用には適さない。また人間に匹敵するような精度が得られている画像データベースは、汎用性を高めるため植物だけでなく人工物や動物など様々な物体が含まれている。すべての種を植物で統一した画像データベースはこれらと比べ種ごとの特徴が乏しくなり、精度の高い識別は困難と考えられ、PlantCLEF 2016 では葉や樹木の枝の画像も含まれているが、1000 種の識別精度は種で 82.2%に留まっている。また植物種数は ILSVRC の 1000 種などと比較して非常に多いため汎用的な画像データベースの構築は難しく、対象地区を絞った範囲での構築が必要と考えられるが、対象地区に生育する種を網羅するような深層学習の応用研究は少ない。

生育する植物種を把握するには調査が必要であるが、調査を伴った研究例として、小型 UAV で得られた画像から植物種の特徴量を抽出し、機械学習によって植物の検出や識別を行う研究がなされている⁷⁾⁸⁾。これらの研究例では機械学習に用いる特徴量を色情報から取得し、その有効性が示されている。しかし色情報の分布が近い種の場合、特徴量の設定が困難であり、識別種を増やした際の運用が困難と思われる。また中根らによる、ドローン空撮画像を矩形グリッドに分割し、各グリッド画像を深層学習法で識別した研究¹⁰⁾では、ドローン空撮や深層学習法の有効性が示されているが、樹木の境界が混合しており、こちらも識別種を増やした際の運用が困難と思われる。境界の混合については鈴木らが MeanShift 法¹¹⁾を用いたスーパーピクセル分割の有効性が示されている。

以上より本研究では対象地区の調査を行い得られた植物画像を基に、深層学習法による植物識別システムの構築、改善を行い、高精度化に必要な撮影手法やデータベースの構築法について検証し、システムの構築化を目指したうえで、都市緑地保全への適用について検討する。

2.2 深層学習法を用いた鳥類検出に関する既往研究

鳥類種をはじめ野生動物のカメラ観察は定常性において有用である。鳥類モニタリングに関する情報技術の応用研究としては、ステレオ計測を用いた飛翔中の鳥類の三次元軌跡計測システムの開発¹²⁾や、滑走路周囲 1km 以下の範囲の鳥類の検出を目的として、自然背景光でも適用できる動体検出法とステレオ PTV を組み合わせた、飛翔体モニタリングシステムの開発¹³⁾、等がある。これらは同時に撮影された複数の動画から、背景差分法を用いて自動検出した鳥類に対し、ステレオの原理から鳥類の空間座標を導出している。しかし背景差分法では動体を機械的に検出してしまうので、動体の中から鳥類を識別する必要がある。またカメラ観察にて得られた膨大な映像データを人手で詳細に解析することは困難であり、動体の鳥類識別と合わせて機械学習、特に深層学習を応用した画像解析の自動化研究が活発に行われている。風車とのバードストライクを対象に、FCN (Fully Convolutional Network)、CNN (Convolutional Neural Network)、SVM (Support Vector Machine) など複数の機械学習法を用いて、風車付近の鳥類を検出する

システムの開発¹⁴⁾や風車付近へ飛来するタカとカラスを識別するシステムの開発¹⁵⁾が行われ、また鳥類の生態系調査のため、ドローンなどの小型無人航空機を用いた広域撮影画像から群れの各個体を検出することも試みられている^{16),17)}。動画解析においてはYOLO (You Only Look Onse)¹⁸⁾で鳥類を検出し、ResNet-50で識別するシステムの開発¹⁹⁾、独自AIカメラによる長期間定点撮影とFasterR-CNN (Faster Regions with Convolutional Neural Network)を用いた、シラサギと他の鳥類の識別とシラサギの大小の生育分布の解析²⁰⁾、風車周りの遠景画像に写る微小な鳥類の検出²¹⁾や、IOT機器とYOLOを用いた哺乳類と鳥類の検出²²⁾や、撮影画像とそれを基にした3Dモデルによる学習データ増強手法を用いたYOLOの害鳥認識システムの開発²³⁾が行われている。各手法の性能比較では、YOLOとFasterR-CNN、R-FCN (Region-based Convolutional Neural Network)およびSSD (Signal Shot MultiBox Detector)の比較では、精度はFasterR-CNN、処理速度はYOLOが優位であると報告がある¹⁷⁾。データサイエンスコンペディションプラットフォームKaggle²⁴⁾の鳥類データセットを用いた鳥類の識別ではYOLOv5のYOLO-v5xモデルがFasterR-CNNなど他の手法と比べ最も高精度となった²⁵⁾。これらの中でYOLOは高速性と高精度を両立した手法であることが伺え、リアルタイム検出へ活用できる性能を持っているが、YOLOを動画の飛跡解析に応用した先行研究は確認できない。カメラによる定点観測では、映像内サイズが非常に小さい遠方を飛翔する鳥類が撮影されるケースも多く、このような鳥類の検出は容易ではないと考えられるが、鳥類の検出最小サイズなどYOLO利用の限界に関する情報は少ない。また、深層学習法による映像解析は有用であるが、撮影する方法、場所、対象により背景や物体のスケールが異なるため、適用ケースごとに学習済みモデルの選択や入力パラメータの最適化が不可欠となる。

以上より本研究では響灘ビオトープを対象に、深層学習法YOLOを用いた鳥類のリアルタイム検出について、適切なパラメータ設定と学習済みモデルの選択を検証し、鳥類のリアルタイム自動検出システムを構築したうえでシステムの常時運用を行う。常時運用により得られたデータからシステムの課題、YOLO利用の限界、飛跡解析への利用について検証する。

参考文献

- 1) Sharada, P. M., David P. H. and Marcel, S. (2016): Using Deep Learning for Image-Based Plant Disease Detection, *Front. Plant Sci.*, Vol.7, pp.1-10
- 2) Wang, Z., Hu, M. and Zhai, G. (2018): Application of Deep Learning Architectures for Accurate and Rapid Detection of Internal Mechanical Damage of Blueberry Using Hyperspectral Transmittance Data, *Sensors*, vol.18, p.1126
- 3) Hasan, M.M, Chopin, J.P., Laga, H. and Miklavcic, S.J. (2018): Detection and analysis of wheat spikes using Convolutional Neural Networks, *Plant Methods*, vol. 14, p.100
- 4) ハナノナ, <https://www.it-chiba.ac.jp/skytree/exhibitions/%E3%83%8F%E3%83%8A%E3%83%8E%E3%83%8A/>, 参照 2022-12-09
- 5) PlantCLEF2016, <https://www.imageclef.org/lifeclef/2016/plant>, 参照 2023-02-03
- 6) Marco S., Michael R., David B., Jana, W. and Patrick, M. (2019): Image-based classification of plant genus and family for trained and untrained plant species, *BMC Bioinform.* Vol.20, No.4
- 7) 山本裕加・田中健太(2019): ドローンによる草索性植物の花の検出, *つくば生物ジャーナル*, Vol.18, p.89
- 8) 小熊宏之・井手玲子・井鷲裕司(2016): UAV 観測画像を用いた絶滅危惧植物の花の自動検出手

法, Journal of Remote Sensing Society of Japan, Vol.36, pp.72-80

9) 鈴木太郎, 土屋武司, 鈴木真二, 山場淳史(2016): 小型 UAV による空撮画像からのスーパーピクセル分割と機械学習を用いた植生分類手法の構築, Journal of The Remote Sensing Society of Japan, vol. 36, pp.59-71

10) 中根英昭・若槻祐貴: 環境分野への深層学習応用研究の立ち上げについて, 高知工科大学紀要, 15(1), pp. 111-120, 2018

11) D. Comanicu and P. Meer, (2002) : Mean Shift: a Robust Approach Toward Feature Space Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, No.5, pp.603-619

12) 神田哲志・村井祐一・田坂裕司・武田靖・手塚英昭・森治嗣 (2009) : ステレオ撮影を用いた鳥の三次元軌跡計測システムの開発, 日本機械学会論文集(B 編), Vol.75, No.751, pp.175-177

13) 村井祐一・田坂裕司・武田靖 (2009) : 応用ステレオ PTV による大気中の飛翔体の 3 次元計測, 可視化情報学会誌, Vol.29, No.1, pp.371-372

14) Takeki, A., Trinh, T.T., Yoshihashi, R., Kawakami, R., Iida, M. and Naemura, T. (2016): Combining deep features for object detection at various scales: finding small birds in landscape images, 情報処理学会 Transactions on Computer Vision and Applications, Vol.8, No.1, pp.1-7

15) Yoshihashi, R., Kawakami, R., Iida, M. and Naemura, T. (2017): Bird detection and species classification with time - lapse images around a wind farm: Dataset construction and evaluation, Wind Energy, Vol.20, No.12, pp.1983-1995

16) Ben, G. W, Lindsey, G., Vienna, R. S, Ashley, S., Andrew, O., Kristen, B., Glenda, Y., Ann, E. M., Rowan, C., Christopher, D. L., Alex, W., Nick, D. H., Alice, J. E., Tom, H., Mark, J. J., Rohan, H. C., Dominik, M., Henry, S., Ryan, D., Ethan, P. W., Peter, F. and S. K. Morgan, E. (2021): A general deep learning model for bird detection in high resolution airborne imagery, ECOLOGICAL APPLICATIONS, <https://www.biorxiv.org/content/10.1101/2021.08.05.455311v2>, 参照 2022-11-10

17) Hong, S.-J., Han, Y., Kim, S.-Y., Lee, A.-Y. and Kim, G. (2019): Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery, Sensors, Vol.19, No.7, pp.1651-1666

18) Joseph, R., Santosh, D., Ross, G. and Ali, F. (2016): You Only Look Once: Unified, Real-Time Object Detection, Computer Vision and Pattern Recognition Conference, 2016, pp.779-788

19) Sarah, H. (2019): A deep learning pipeline for wildlife camera data analysis, CS230 Deep Learning, Autumn 2019, Stanford University, https://cs230.stanford.edu/projects_fall_2019/reports/26261732.pdf, 参照 2022-11-10

20) Mao, X., Chow, J.K., Tan, P.S, Liu, K.-F., Wu, J., Su, Z., Cheong, Y.H., Ooi, G.L., Pang, C.C. and Wang, Y.-H. (2021): Domain randomization-enhanced deep learning models for bird detection, Scientific Reports, Vol.11, No.1, 639

21) Alqaysi, H., Igor, F., Faisal, Z. Q. and Mattias, O'N. (2021): A Temporal Boosted YOLO-Based Model for Birds Detection around Wind Farms, Journal of Imaging, Vol.7, No.11, pp.227-239

22) Shim, K., Andre, B., Napoleon, R., and Nasim, A. (2021): Small mammals and bird detection using IoT devices, 36th International Conference on Image and Vision Computing New Zealand, <https://ieeexplore.ieee.org/document/9653430>, 参照 2022-11-10

- 23) 北風裕教・吉原蓮人・岡部蒼太・松村遼 (2020): オブジェクト検出 YOLO を用いた害鳥認識システムの開発, 産業応用工学会論文誌, Vol.8, No1, pp.10-16
- 24) Kaggle, <https://www.kaggle.com/>, 参照 2022-11-10
- 25) Zhihao, L. and Ming, T. (2021): A bird species detection method based on YOLO-v5, 2021 International Conference on Neural Networks Information and Communication Engineering, <https://doi.org/10.1117/12.2615310> , 参照 2022-11-10

第3章 研究手法

3.1 畳み込みニューラルネットワーク (CNN) の概要と神経回路網モデルの性能

CNNは画像の特徴抽出を自動で行う技術である。深層学習法以前の機械学習では、識別に効果のある特徴は専門家が経験的に手作業で設計していた、例えば、アヤメには3つの種類があり、その分類は花びらと萼片の長さや幅という専門家が図った4つの特徴量によって決まる。一方でCNNでは特徴抽出を図3-1に示すように自動で行うことができる。CNNにより、人の判断を介さずに画像の分類が行えるようになった。

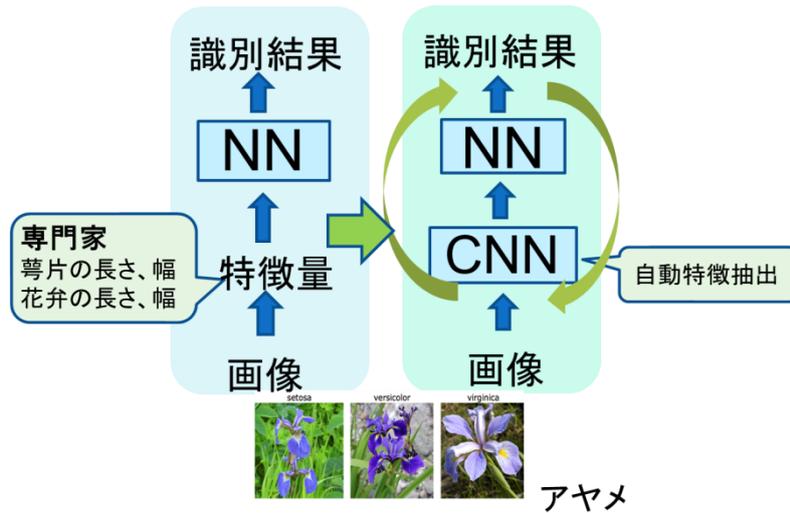


図 3-1 アヤメにおける識別過程

CNNの構成について述べる。CNNは畳み込み層とプーリング層から成り立っている。まず、畳み込み層 (Convolution 層) で入力画像にフィルターをセル各々に対して左上からスライドするように掛け合わせていき、その合計値の行列データを作成する。次に、プーリング層で合計値の行列データのある領域で区切り、それぞれの領域の最大値をその領域の値とし、データを圧縮する。これを繰り返し行うことで識別の誤差を最小にするように、ニューラルネットのつながりの最適値を求めていく。また、畳み込み層のフィルターの値も、学習を反映し学習毎に更新されていく。図3-2～図3-4に各段階における処理のイメージ図を示す。

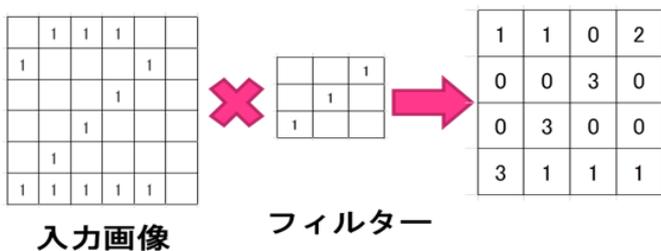


図 3-2 畳み込みのイメージ図

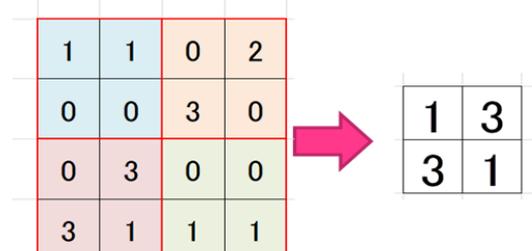
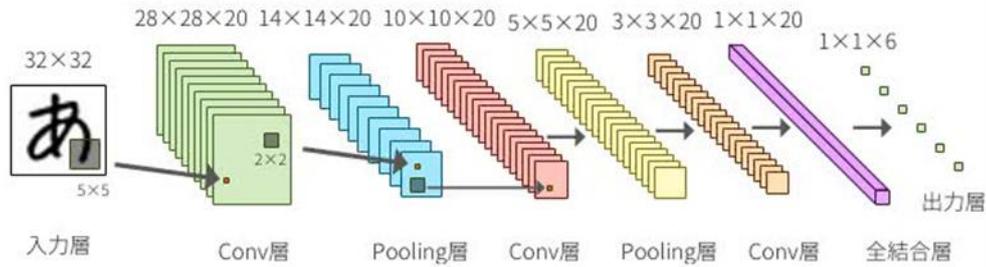


図 3-3 プーリングのイメージ図



引用) https://deepage.net/deep_learning/2016/11/07/convolutional_neural_network.html

図 3-4 CNN と NN のつながり方

3.2 Keras による深層学習法の実装

本研究では第 4 章, 第 5 章で用いた深層学習システムの実装に Keras¹⁾を用いた。Keras は python を開発言語としたニューラルネットワークライブラリで, API を通して多くの神経回路網モデルを扱うことができる。バックエンドエンジンとして, Tensorflow²⁾や Theano, CNTK といった複数の機械学習フレームワークを備えており, 神経回路網モデルによる演算処理は, バックエンドで稼働しているフレームワークが行う。図 3-5 に Keras を用いた深層学習システムの構成図を示す。ユーザは Jupyter notebook などの python 開発環境を介して Keras を扱う。Keras に用意されている神経回路網モデルを利用した演算処理は, 機械学習フレームワークが深層学習に最適化された演算ライブラリ (cuDNN など) を用いて行う。表 3-1 に Keras に備わる神経回路網モデルとその性能を示す。表 3-1 において top1 は予測確率 1 位のみの確度 (Accuracy), top5 は予測確率 5 位までに正答が含まれている場合の確度であるが, これらの確度は ImageNet による画像分類テストによって得られたものである。パラメータ数は層と層をつなぐ重み数の合計値で, 層数はモデルを構成する層数の合計値である。本研究ではバックエンドエンジンとして TensorFlow を起用した Keras を用い, Keras に備わっているファインチューニングやデータオーギュメンテーションを利用して Xception による深層学習システムを構築した。

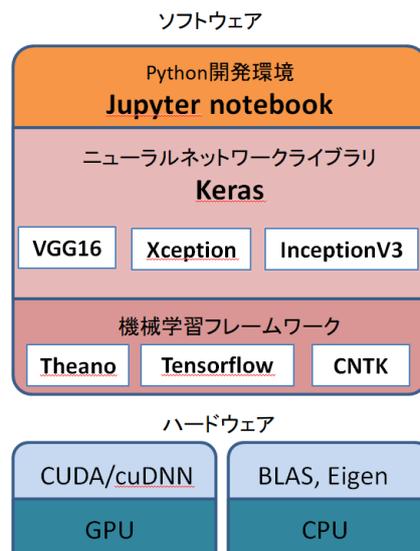


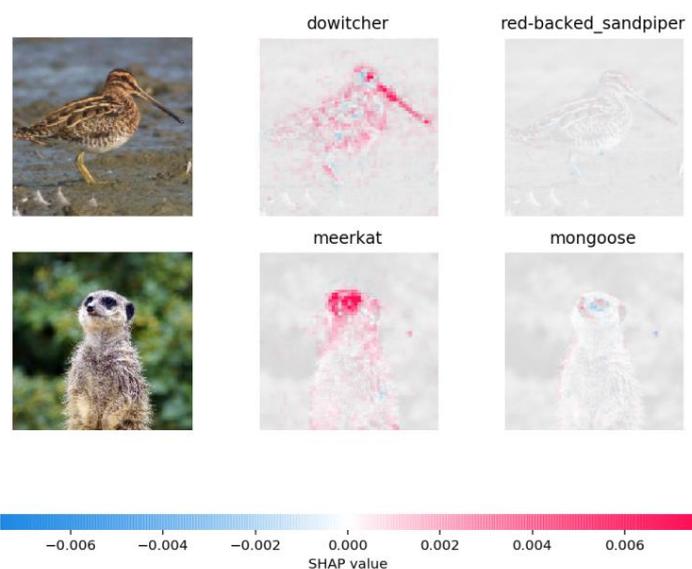
図 3-5 Keras を用いた深層学習システム構成図

表 3-1 神経回路網モデルの性能

モデル名	Top1	Top5	パラメータ数	レイヤー数
Xception	79.0	94.5	22,910,480	126
VGG16	71.5	90.1	138,357,544	23
VGG19	72.7	91.0	143,667,240	26
ResNet50	75.9	92.9	25,636,712	168
InceptionV3	78.8	94.4	23,851,784	159
InceptionResNetV2	80.4	95.3	55,873,736	572
MobileNet	66.5	87.1	4,253,864	88
DenseNet121	74.5	91.8	8,062,504	121
DenseNet169	75.9	92.8	14,307,880	169
DenseNet201	77.0	93.3	20,242,984	201

3.3 Explainable AI による着目領域の可視化

Explainable AI は、深層学習により得られた予測確率の根拠を可視化する手法である。本研究では代表的な手法である SHAP (SHapley Additive exPlanations) 法³⁾を用いて、植物画像識別時の着目領域の解析を行った。SHAP 法は学習モデルの予測結果に対する各特徴量の寄与を求めるための手法で、協力ゲーム理論の ShapleyValue に着想を得ており、学習モデルの平均的な予測値と、ある特徴量が加わった時の予測値との差分を、その特徴量の寄与としている。ShapleyValue のもつ「全体合理性」、「対称性」、「加法性」、「ナルプレイヤーのゼロ評価」といった性質が、特徴量の寄与が持つべき性質と合致しており、他よりも優れた手法とされている。Python での解析には slundberg が提供している SHAP ライブラリ⁴⁾を使用した。図 3-6 に SHAP 法による解析例を示す。画像から抽出されるそれぞれの特徴量が、画像の識別に与えた寄与の度合いを赤及び青の濃淡で可視化している。



引用) slundberg SHAP ライブラリ⁴⁾

図 3-6 SHAP 解析例

3.4 MeanShift 法を用いたスーパーピクセル分割

草本種や木本種を撮影した画像を群落や樹冠ごとに分割するために、本研究では色情報を用いた MeanShift 法によるクラスタリングを用いて、群落や樹冠単位のスーパーピクセルを作成した⁵⁾。MeanShift 法はデータ群の重心を探索することでクラスタリングを行うが、重心を周辺色情報の平均値に置き換えることで、1画素とその近傍が平滑化されていき、近しい色情報を持つ領域が1つのスーパーピクセルとなる。画像処理ライブラリの OpenCV には MeanShift 法を応用した画像の領域分割を行うメソッド、`pyrMeanShiftFiltering()`が実装されている。本研究ではこれを用いて群落や樹冠をスーパーピクセル化し、画像処理ライブラリ Scikit-image に実装されている `slic()`メソッドを用いてスーパーピクセル単位の領域分割を行った。図 3-7 に MeanShift 法によるスーパーピクセル作成の例を、図 3-8 にスーパーピクセル化した画像の領域分割結果を元画像に反映させた例を示す。



図 3-7 元画像(左)と MeanShift 法による平滑化処理を施した画像(右)



図 3-8 スーパーピクセルを基にした領域分割

3.5 YOLO 法による物体検出

YOLO 法⁶⁾は入力された画像サイズを修正し CNN に入力、その後ネットワークによって解析された結果を出力する。これらの入力画像サイズの修正、CNN への入力および解析、結果の出力を1つのフレームワークで行うことができる。出力は検出された一つの物体につき、座標値 (x,y,w,h) と信頼度の計 5 つの値が出力される。図 3-9 に物体検出の概要図を示す。入力画像を縦横にグリッド分割を行う。グリッド毎に周辺にある物体を 2 つ推定する。同時にグリッド毎にそのグリッド領域に映る物体を識別する。すべてのグリッドの走査が終了した後推定物と識別物を統合し物体の検出を行う。最後に同一物の中か

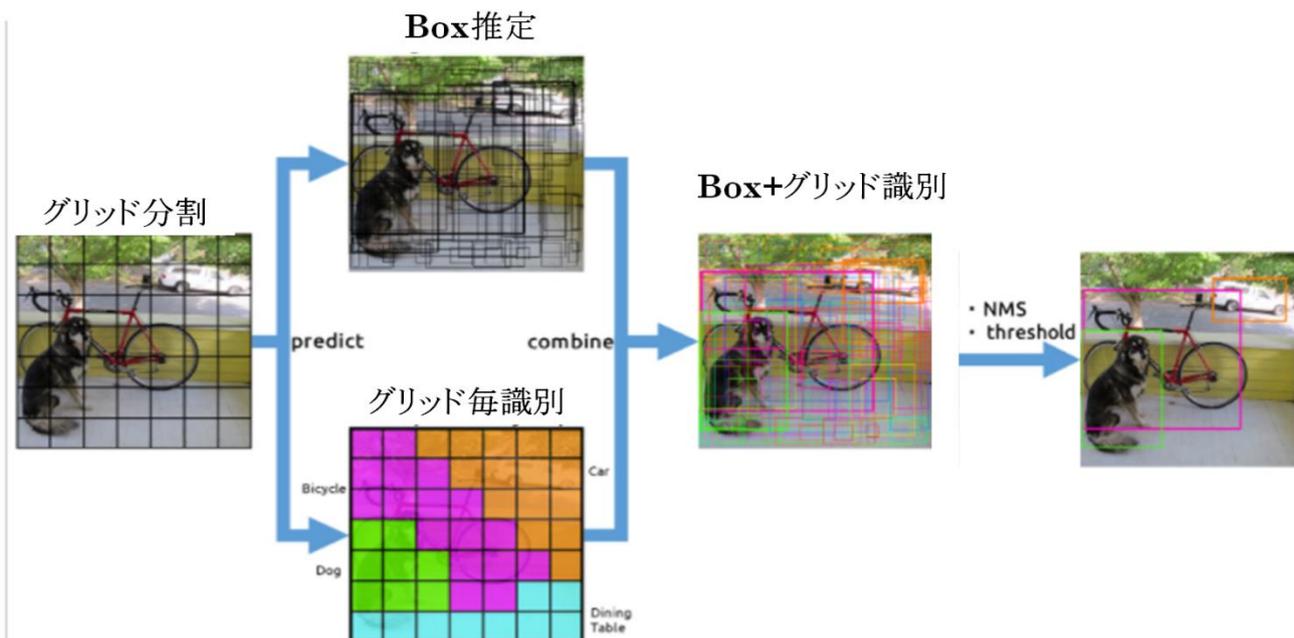


図 3-9 YOLO 法における物体検出手法の概要図

ら一番大きな領域を抽出し検出物とする。

YOLO 法では精度を示す値として mAP (mean Average Precision) を用いている。これは識別種すべての AP (Average Precision) の平均値で、AP は Precision (適合率) と Recall (再現率) の積分によって求められ、それぞれの算出方法を以下に示す。物体検出では、BoudingBox (以下 bbox) という境界線で物体を検出する。この時予測された PredictedBoudingBox (以下 pbbox) と真値 Ground-truth Bounding Box (以下 gtbbox) の誤差を IoU (Intersection over Union) で表す。IoU は 0~1 の値をとり、0 のとき pbbox と gtbbox が全く重なっていない状態、IoU が 1 のとき完全に重なっている状態を指す。ここで IoU が 0.5 以上の時を True、0.5 未満の時を False としたとき、物体検出では以下の 4 つの区分に分けることができる。

- TruePositive(以下 TP) : 物体 X を X と予測した場合
- FalsePositive (以下 FP) : 物体 X でない物体を X と予測した場合
- FalseNegative(以下 FN) : 物体 X を X でないと予測した場合
- TrueNegative(以下 TN) : 物体 X でない物体を X でないと予測した場合

以上の 4 つの指標をもとに Precision と Recall を求める。それぞれの式は以下になる。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

これらの 2 つの値から AP を算出することができ、Precision を p, Recall を r とすると AP は以下のようにならわすことができる。

$$\text{AP} = \int_0^1 p(r) dr$$

参考文献

- 1) Keras, <https://keras.io/>, 参照 2023-02-03
- 2) TensorFlow, <https://www.tensorflow.org/?hl=ja>, 参照 2023-02-03
- 3) Erik, Š. And Igor K. (2014) : Explaining Prediction Models and Individual Predictions with Feature Contributions, Knowledge and Information Systems, vol.41, pp.647-665
- 4) SHAP ライブラリ, <https://github.com/slundberg/shap>, 参照 2023-02-03
- 5) D. Comanicu and P. Meer, (2002) : Mean Shift: a Robust Approach Toward Feature Space Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, No.5, pp.603-619
- 6) Joseph, R., Santosh, D., Ross, G. and Ali, F. (2016): You Only Look Once: Unified, Real-Time Object Detection, Computer Vision and Pattern Recognition Conference, 2016, pp.779-788

第4章 深層学習法による植物画像識別の植物調査への応用

4.1 はじめに

生態系のベースとなる植物種の把握や、その分布を示す植生図の作成は、都市部における河川敷や公園等緑地の保全活動において重要であるが、そのための調査が十分になされている地区は限定されている。福岡県北九州市では、板櫃川が毎年水辺の国勢調査を行っているが、同じく植生を伴う河川改修へ転換された撥川や、重要種が保全されているひびきの北公園、響灘ビオトープでの生態系調査は不定期でしか行われていない。これは希少動物種の保存にとって重要な植物生態系の重要性が十分に認識されていないことに加え、植物種の識別と撮影に専門家や経験者が必須となり、調査の頻度や地域が制限されるためである。これら保全上の課題は全国の都市緑地の多くに共通したものであり、保全上重要なモニタリングの困難さを軽減する植物種の効率的な分類手法が必要とされている。北九州市のこれらの緑地、ビオトープ等は都市緑地に共通する属性を持った典型的な緑地であるために、これらの地区における植物種の分類の効率化を行うための手法について研究することは、全国的にも意義がある。そのため高い認識精度が得られている深層学習法の画像認識技術を用いることは有効と考えられる。

本章の予備的研究として、ImageNet¹⁾に属する47種の花画像を用い、植物種の分類に適した神経回路網モデルの評価を行った。Xception²⁾モデルでハイパーパラメータを最適化することによって、90%以上の確度(Accuracy)を達成した³⁾。またImageNetには間違った分類をされた画像や、低解像度で種の判別が困難な画像も含まれ質に問題があるため、植物園で独自に撮影した53種の植物画像を用いた学習モデルを作成した。枝葉等の細部情報を補完するため、切り抜き画像を補完に用いた画像と合わせて学習させることにより、Xceptionモデルで95%以上の確度を得た。以上より、これらの手法を用いて都市における保全緑地の植物種の分類を高い確度で効率的に行うためには、調査緑地の数や季節毎の調査を継続し、植物種の質の高い画像データベースと学習モデルを構築することが重要であると考えられる。これを受けて2019年1年間の調査では、2地域で81種の植物を撮影し、作成した学習モデルは95%以上の確度を得た。しかし調査地点とは別の地点で撮影した同種の植物の識別を行うと、植物種によって確度のバラツキが大きく、6割の植物種は80%以上であるものの3種は20%以下であり、これらの改善が課題となった。⁴⁾

本章では、①植物種によって確度が大きく低下する要因を解明して植物の撮影や学習方法の改善指針を得ること、②改善指針に基づき新たな植物調査と学習を行い改善の効果を確かめること、③新たな植物調査と学習を行い未知データに対する識別で確度を向上させて汎化性能を高めることを目的とし、継続的に植物調査を行い学習モデルの作成と改善を行った。

4.2 植物種画像データベースと学習モデルの構築手法

4.2.1 植物調査地域および調査の概要

ひびきの北公園(北九州市若松区塩屋2丁目)、響灘ビオトープ⁵⁾(北九州市若松区響町1丁目)、撥川河川敷⁶⁾(北九州市八幡西区岸の浦)の3地区を調査対象とし、専門家同伴の下で植物種を調査した。植物調査の調査方法及び各地区の調査日を表4-1に示す。なお、ひびきの北公園及び撥川河川敷におけるライントランセクトは2m幅で行った。ひびきの北公園は用地の一部を北九州学術研究都市の里山保全地区として整備しており、キンラン、ハンノキなどの希少種を移植、保全しているが、保全地区北部の外縁を調査区とした(図4-1)。響灘ビオトープは41haの大きさを有し、湿地や淡水池、草原等の多様な環境

表 4-1 植物調査の概要

調査地点	ひびきの北公園	響灘ビオトープ	撥川河川敷
調査方法	ライントランセクト法 ⁷⁾	コドラート法 ⁸⁾	ライントランセクト法
調査日	2019/5/9	2019/6/11	2020/5/19
	2019/9/12	2019/10/9	2020/10/1
	2020/5/21	2020/5/26	
	2020/9/23	2020/10/23	



図 4-1 ひびきの北公園での調査地点



図 4-2 響灘ビオトープ⁵⁾での調査地点



図 4-3 撥川での調査地点

であることから 100 平米のコドラートを設け、湿地 3 地点および台地 3 地点の計 6 地点で調査を行った。響灘ビオトープの調査地点を図 4-2 に示す。調査時期は開花期及び結実期等の種の同定が行いやすいため、春から初夏と秋を含む年 2 回を設定した⁹⁾。撥川河川敷は河川公園の周辺を調査地点とした (図 4-3)。

4.2.2 深層学習方法

深層学習フレームワークには Keras¹⁰⁾/Tensorflow を、神経回路モデルには Xception²⁾ を用いた。深層学習のハイパーパラメータは予備的研究にて最適化した値を用い、エポック数を 30、バッチサイズを 16 に設定し、学習率は 0.01 から 5 エポック毎に減衰率 0.6 で小さくするよう設定した。学習はカラー画像を用いて行い過学習を防ぐためのデータ増強処理は Keras が標準で用意している設定を用いた。

深層学習に用いた画像の撮影には Panasonic 製デジタルカメラ DC-FZ85 を用い、解像度 4896×3672 のオリジナル画像を取得した。枝葉等の細部情報も学習させるため、オリジナル画像から 1024×1024, 2048×2048 の画像を切り出した (以下、部分画像と記す)。切り出す範囲は乱数を用いて無作為に決定し、画像間の重なりも考慮して最大 10 枚ずつの部分画像を作成した。部分画像を加えた全画像データを学習、検証、試験それぞれに 7:2:1 の割合でランダムに振り分けた。目標識別精度については表 3-1 より学習に用いた Xception の Top1 確度と Top5 確度の平均値 86.7% を種ごとの識別において Top3 確度で目指す。

4.2.3 XAI 解析

植物種により確度が低下する要因を解明するため、神経回路網モデルが画像識別時に着目する領域を可視化できる XAI (Explainable AI) を導入した。XAI の代表的な手法の 1 つである SHAP 法¹¹⁾を用いて植物画像識別の判断部分の解析を行った。SHAP 法により画像から抽出される特徴量が画像識別の際に与えた影響を識別画像に重ねて可視化した。

4.3 調査・学習結果

4.3.1 2019 年 2 調査地点の深層学習結果

2019 年に調査したひびきの北公園 (春期・秋期) 30 種、響灘ビオトープ (夏期・秋期) 57 種、2 地点を総合した 81 種 (2019 モデル) の深層学習結果を表 4-2 に示す。学習時の確度は学習中に検証用画像で識別した際の確度であり、試験時の確度は学習後に学習に用いていない試験画像を識別した際の確度である。調査地点別の試験確度は約 99%、総合した 81 種の試験確度も約 98% であり、植物種が 81 種まで増えても確度が維持されていることが分かる。

この極めて高い確度を与える深層学習結果は非常に有望であるものの、撮影地点が狭くオリジナル画像の撮影枚数も少ないため、その一因として画像に映り込む背景の影響が考えられる。そのため、響灘ビオトープで 2019 年秋期に調査とは異なる別地点で植物種を撮影した。撮影した画像 14 種 107 枚の分類結果を表 4-3 に示す。予測確率 1 位を 1st, 2 位を 2nd, 3 位を 3rd とし、top1 は予測確率 1 位のみの確度、top3 は予測確率 3 位までに正解が含まれている場合の確度を示す。高い確度が得られた種はアキノノゲシ (図 4-4)、アメリカウンランモドキ (図 4-5)、スギナ (図 4-6) で 100%、アレチハナガサ、ヘクソカズラ、ヒメジョオン、セイタカアワダチソウ、ヨシで 80% 以上だった。確度が低い植物はクズ (図 4-7) が 20%、チガヤ (図 4-8) とタチスズメノヒエ (図 4-9) は 0% であった。確度が高い植物はアキノノゲシ、アメリカウンランモドキなど花が含まれるか、スギナなど特徴的な茎や葉が画像全体を占めている

表 4-2 2019 年の深層学習結果

	植物種	撮影枚数	確度	
			学習時	試験時
ひびきの北公園	30	1242	99.6	99.1
響灘ビオトープ	57	2368	96.4	98.8
総合	81	3610	97.9	97.9

表 4-3 別地点で撮影した画像による分類試験結果

種類	画像数	正解画像数			確度	
		1st	2nd	3rd	top1	top3
アキノノゲシ	2	2	0	0	100.0	100.0
アメリカウンランモドキ	16	16	0	0	100.0	100.0
アレチハナガサ	12	10	0	1	83.3	91.7
チガヤ	3	0	0	2	0.0	66.7
ヘクソカズラ	10	8	0	0	80.0	80.0
ヘラオオバコ	2	1	0	0	50.0	50.0
ヒメジョオン	9	8	1	0	88.9	100.0
クズ	10	2	1	1	20.0	40.0
セイバンモロコシ	5	3	0	0	60.0	60.0
セイタカアワダチソウ	7	6	0	1	85.7	100.0
スギナ	6	6	0	0	100.0	100.0
タチスズメノヒエ	6	0	1	0	0.0	16.7
ヨモギ	12	5	2	1	41.7	66.7
ヨシ	7	6	1	0	85.7	100.0



図 4-4 アキノノゲシ



図 4-5 アメリカウンランモドキ



図 4-6 スギナ



図 4-7 クズ



図 4-8 チガヤ



図 4-9 タチスズメノヒエ

ような種が多い傾向にあった。

確度が低い 3 種は背景や他の植物の映り込みの影響が考えられるため、細部を切り出した部分画像 (1024×1024 もしくは 2048×2048) を用いた識別を行った。結果を表 4-4 に示す。確度は向上し、チガヤやクズは 45%, 42% となった。一方で、122 枚の画像は対象植物が中央に写っているにもかかわらず top3 でも識別できなかった。これらクズ・チガヤ・タチスズメノヒエの 3 種について、確度が低下する要因を詳細に解析した。

表 4-4 部分画像を含めた分類試験結果

種類	画像数	正解画像数			確度	
		1st	2nd	3rd	top1	top3
チガヤ	42	19	7	6	45.2	76.2
クズ	227	96	30	18	42.3	63.4
タチスズメノヒエ	62	9	8	5	14.5	35.5

4.3.2 2019 年学習結果における確度の低い植物の要因解析

(1) クズ

学習に用いたクズの画像 1094 枚を小葉の枚数別に分類した。大きく一枚の小葉を映しているものや完全に小葉 1 枚が写りきっていないものを「1 枚以下」、それ以上を「2~9 枚」、「10 枚以上」と分類した結果、学習画像の 39.8% が「1 枚以下」、54.6% が「2~9 枚」、5.6% が「10 枚以上」であった。それに対し、別地点の試験画像は「10 枚以上」のものばかりであり、学習用画像との小葉の枚数の違いが確度に影響している可能性がある。学習用画像の小葉の枚数を増やすため、切り出す部分画像のサイズを 1024×1024, 2048×2048 から 2048×2048, 3072×3072 に拡大して学習させた結果を表 4-5 に示す。別地点画像（全体画像）の確度は改善しないものの部分画像は top3 までなら 81% まで向上した。また部分画像を完全に写っている小葉の枚数で分けし、枚数別にクズの予測確率(probability)の平均値を算出した。解析結果を図 4-10 に示すが、小葉 2 枚で確度は 60%, 3 枚で 71% となった。予測確率の平均値は 2 枚の 0.42 から、3 枚では 0.75 まで向上した。クズは、3 小葉で 1 つの葉である¹²⁾ ことから、クズの確度向上には、3 枚の小葉が完全に写っている画像を用いることが望ましいと考えられる。

表 4-5 クズの学習時の小葉枚数を増やした試験結果

	画像数	正答画像数			確度	
		1st	2nd	3rd	top1	top3
学習後試験	201	201	0	0	100.0	100.0
別地点画像	10	1	2	0	10.0	30.0
部分画像	217	113	42	22	52.1	81.6

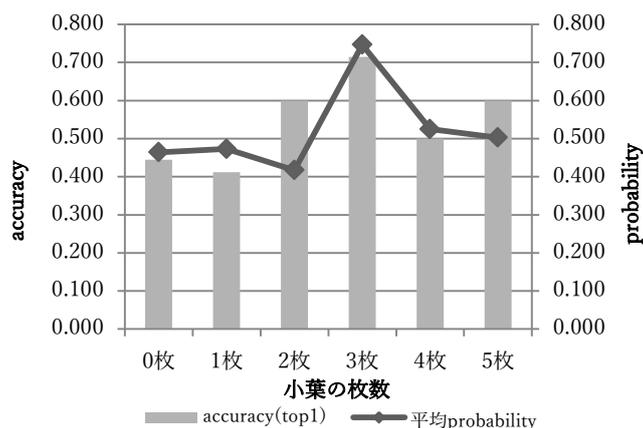


図 4-10 小葉の枚数ごとの確度と平均予測確率

(2) チガヤ

表 4-4 よりチガヤの部分画像の確度は 45%に向上したが、別地点で撮影した試験画像には穂が映っており、穂と葉の識別感度を比較するため、1024×1024 の部分画像を葉のみの画像（図 4-11 左）と、穂も写っている画像（図 4-11 右）に分けて識別テストを行った。各部分画像の結果を表 4-6 に示す。正方形サイズ 1024 画像で、葉のみの確度が 18%に対して穂が写っている画像では確度が 83%まで向上した。オリジナル画像でも穂は写っているが、部分画像で確度が向上する理由を調べるため SHAP 解析を行った。

オリジナル画像の結果を図 4-12、部分画像の結果を図 4-13 に示す。図 4-12 より、原画像でも穂に注目しているがその領域は小さく、約 98%の予測確率でオギと認識しているのに対し、チガヤの予測確率は約 5%に止まっていた。図 4-13 の部分画像では拡大された穂全体に反応し、チガヤを最も特徴づける穂の形態が明確になったことで、約 80%の予測確率でチガヤと認識できたと考えられる。Xception の入力サイズが 299×299 であり、これより大きな画像は入力時に縮小されることも考慮すると、学習画像は被写体の特徴を大きく捉える必要があると考えられる。また、葉のみの場合でも同じイネ科のセイバンモロコシ（別地点テストの確度が 60%）を top3 で解答していることから、植物種までは識別できていないがイネ科に共通の特徴を認識できていると思われる。他の誤答はウキヤガラやヒメガマなどで、学習画像において無造作に葉が生えているような種に識別したと考察できる。



図 4-11 チガヤ学習画像例（左穂なし，右穂あり）

表 4-6 チガヤにおける試験結果

正方形サイズ(枚数)	top1	top3
3672(3)	0.0	66.7
2048(32)	42.9	59.5
1024(38)	30.8	53.8
1024 葉(32)	18.8	43.8
1024 穂(6)	83.3	100.0

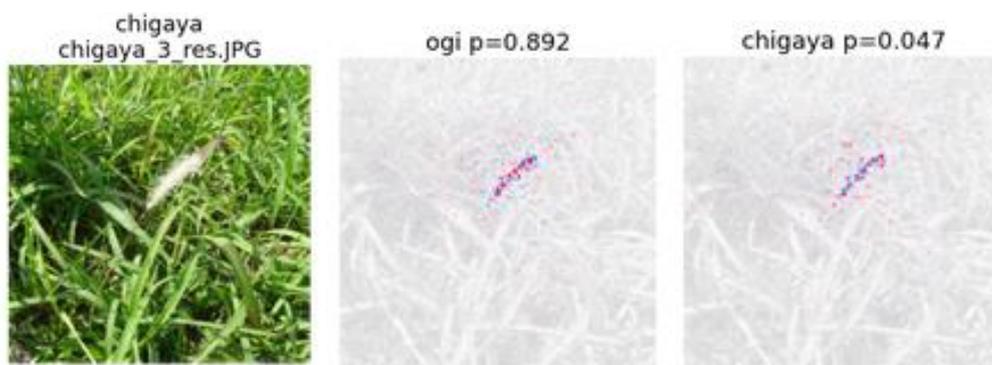


図 4-12 チガヤのオリジナル画像を用いた SHAP 解析の結果

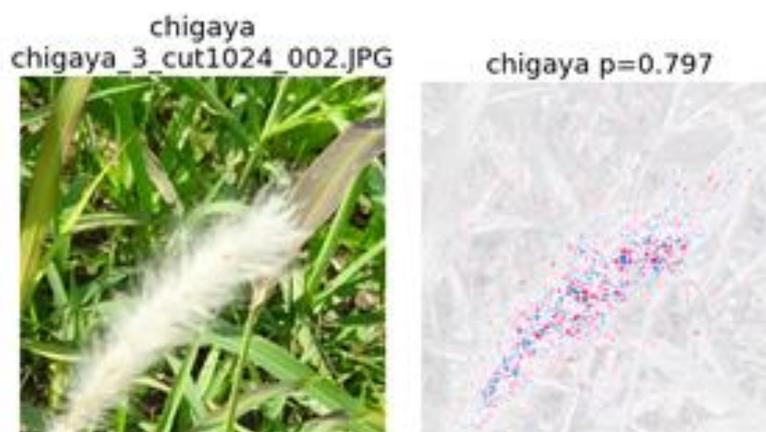


図 4-13 チガヤの部分画像による SHAP 解析結果

(3) タチスズメノヒエ

タチスズメノヒエは、穂が細く色も背景と似たような緑色であるため、背景の影響を大きく受けると考えられる。背景の影響を排するため、ポリゴン分割によって対象植物だけを切り抜いて識別テストを行った(図 4-14)。結果オリジナル画像と切り抜き画像での差異はなく、タチスズメノヒエの予測確率は1%以下だった。図 4-15、図 4-16 に SHAP 解析画像を示す。図 4-15 はチガヤと識別したが、画像中の識別に反応した領域は背景のロープのみであり、これをチガヤの穂と誤認したと推測できる。図 4-15 を切り抜いた図 4-16 では、背景の除去が不十分であるが、対象の植物に反応していることが確認できる。図 4-14 の SHAP 解析では切り抜き画像において穂の部分に反応し、メヒシバに次ぐ2番目の確率でタチス



図 4-14 タチスズメノヒエ のオリジナル画像 (左) とその切り抜き (右)



図 4-15 タチスズメノヒエの SHAP 解析画像

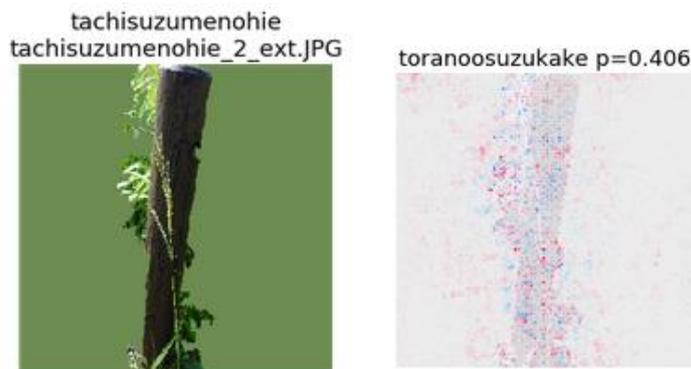


図 4-16 切り取り画像の SHAP 解析

ズメノヒエを予測した。

4.3.3 2020 年春季調査を追加した 2020 春モデルの深層学習結果

2020 年春季調査は撥川を含む 3 地点で行ったが、2019 年で得られた要因解析結果に基づき、被写界深度を狭くし、穂や茎が細い種は後ろに板を置いて対象植物以外の背景の影響を小さくなるように撮影した。また学習に用いる部分画像は対象植物が中心に写っている画像のみを選別した。2020 年春季の調査にお

いて撮影した植物種の一覧と学習および学習後テストの結果を表 4-7 に示す。2019 年と 2020 年春季に行った計 7 回の植物調査で撮影した植物は全 103 種となった。これらを 1 つの画像データセットに統合し、部分画像についても対象植物が中心に写っている画像のみを選別して学習したものを 2020 春モデルとした。深層学習結果を表 4-8 に示す。top1 の確度は 90% となったが、これは識別植物が 103 種まで増加したことが原因と考えられる一方で、top3 は 95% 以上の確度で識別していることが確認できる。

2020 春モデルについても表 4-3 と同様に別地点 14 種での識別テストを実施した。14 種の平均は 2019 年のみの学習モデルと大きな差はなかったが、種別にみると確度の低かった 3 種（クズ・チガヤ・タチスズメノヒエ）の確度は表 4-9 に示すように向上した。また 2020 春モデルでは、タチスズメノヒエの切り抜き画像の予測確率が向上し、図 4-14 は 10 位から 2 位、図 4-15 は 6 位から 4 位となった。

一方で表 4-9 に掲示したヘラオオバコ、アレチハナガサ、ヨモギの確度は下がった。ヘラオオバコは試験画像 2 枚とも不正解となったが、学習画像に全体像が少ないことが原因であると考えられる。アレチハナガサは 2020 春モデルでは細部を拡大した画像の割合が増えたため、相対的に背景が多い試験画像

表 4-7 2020 年春季の深層学習結果

	植物種	撮影枚数	確度	
			学習時	試験時
撥川河川敷	20	682	99.9	93.7
ひびきの北公園	21	945	99.9	97.8
響灘ビオトープ	27	800	99.6	92.5
総合	50	2427	99.7	95.5

表 4-8 2020 春モデルの深層学習結果

	確度	
	学習時	試験時
top1	99.0	89.8
top3	99.9	95.5

表 4-9 分類試験結果(14 種のうち 6 種抜粋)

種類	試験画像数	正解画像数			確度	
		1st	2nd	3rd	top1	top3
クズ	10	4	1	3	40.0	80.0
チガヤ	3	3	0	0	100.0	100.0
タチスズメノヒエ	6	1	0	0	16.7	16.7
ヘラオオバコ	2	0	0	0	0.0	0.0
アレチハナガサ	12	8	0	0	66.7	66.7
ヨモギ	12	1	1	1	8.3	25.0

(全体画像)では確度が低下したと考えられる。ヨモギは葉の形が明確に写っていない画像は誤答が多く、葉が画像の全面に明確に写っている場合は確度が高い。ヨモギの花の画像では、ギンギシやアレチギンギシと誤答する傾向があった。これは2020年春の調査では葉の画像は追加されたが、花の画像の追加は無いため2020春モデルでは葉への反応がより強くなった結果と推察できる。

4.3.4 2020 春モデルにおけるヨモギの確度低下の要因解析とその改善

ヨモギは表 4-3 では確度が約 67% (top3) であったのに対し、表 4-9 では約 25% (top3) と大幅に減少したため、試験に用いた 12 枚の画像について詳細解析を行った。これらは写り方に違いがあり、葉が画像内の一部、全面、ヨモギの花画像の 3 タイプに誤答を分類できた。図 4-17 に葉が画像内の一部に写っているヨモギ画像とこの画像を用いた識別結果の予測確率を示す。最も高い予測確率でトゲミノボタンと回答しており、次いで回答しているセリも含め、葉の形が似ていることが誤答の原因と考えられる。また葉が画像の全面を占めている画像では、予測確率 99.9% でヨモギと回答した。図 4-18 にヨモギの花画像とこの画像を用いた識別結果の予測確率を示す。ギンギシ・アレチギンギシで 99% 以上の予測確率を示していることがわかる。

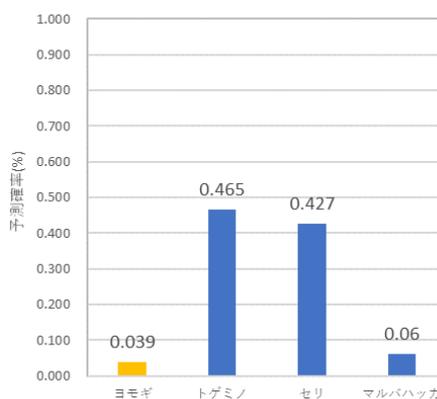


図 4-17 ヨモギの葉画像とその回答例

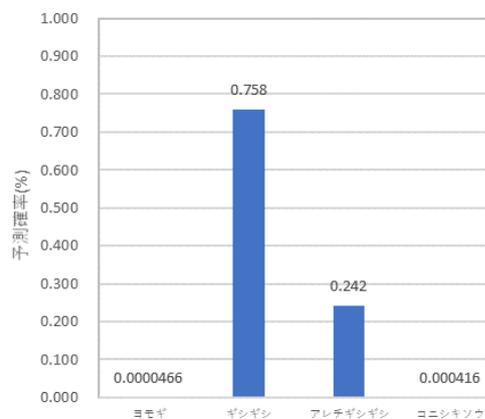


図 4-18 ヨモギの花画像とその解答例

さらに 12 枚の試験画像から部分画像を 97 枚作成し、画像に写るヨモギの葉及び花の占有度による確度の傾向を探った。これらの画像を①葉が画像の全面を占める場合、②葉が画像内の 50%以上を占める場合、③葉が画像内の 50%未満の場合、④花が画像全面に写っている場合、⑤花が画像の一部しか写っていない場合に分類して傾向を探った結果を表 4-10 に示す。画像内に占める葉の面積が多いほど、ヨモギと回答する率が高くなる傾向がわかる。①の場合は 12 枚中 10 枚が top1 でヨモギと回答しており、その予測確率も 93%以上で正答できていた。次いで回答している種としてトゲミノキツネノボタンやセリで誤答している画像が多い傾向があった。②の場合は 27 枚中 15 枚が top1 でヨモギと回答しており、うち 9 枚が予測確率 90%以上であった。また top2 では 23 枚が、top3 では 25 枚がヨモギであると回答した。誤答例としてはトゲミノキツネノボタンやマルバハッカが多かった。③の場合は 20 枚中 4 枚が top1 でヨモギと回答しており、うち 3 枚が予測確率 92%以上であった。ヨモギを 2 番目に回答している画像は 7 枚あったがいずれも予測確率が 20%から 0.1%未満と低く、誤答する種のばらつきも大きかった。①と比較してマルバハッカと回答する確率が 46%、②と比較して 18%高くなっていることから写っている葉の枚数が少なくなるとマルバハッカと回答する傾向があると考えられる。また、④の花画像は 25 枚中 top3 に入ったのは 3 枚で、Top1 で回答している種ではアレチギンギシやギンギシ、ヨシが見られた。⑤の花画像に対しては誤答する種のばらつきが大きく、ヨモギと認識するのは難しいと考えられる。

SHAP 解析を行うと、図 4-19 より葉画像は全面に写っている画像は上に示している画像に比べて拡大した画像になっているため、葉の形を認識しヨモギに分類したと考えられる。また、図 4-20 より花画像は下の全面に写っている方は花に加えて葉も写っているためヨモギに分類できたのではないかと考えられる。

表 4-10 ヨモギ の部分画像のパターン分けによる識別結果

オリジナル画像	部分画像	占有度	試験画像数	正解画像数			誤答例
				1st	2nd	3rd	
葉全面	①	全面	12	10	1	0	トゲミノキツネノボタン、セリ
葉一部	②	50%以上	27	15	8	2	トゲミノキツネノボタン、マルバハッカ
	③	50%未満	20	4	7	1	種のばらつきが大きい
花	④	全面	25	1	0	2	アレチギンギシ、ギンギシ、ヨシ
	⑤	一部	12	0	1	0	種のばらつきが大きい

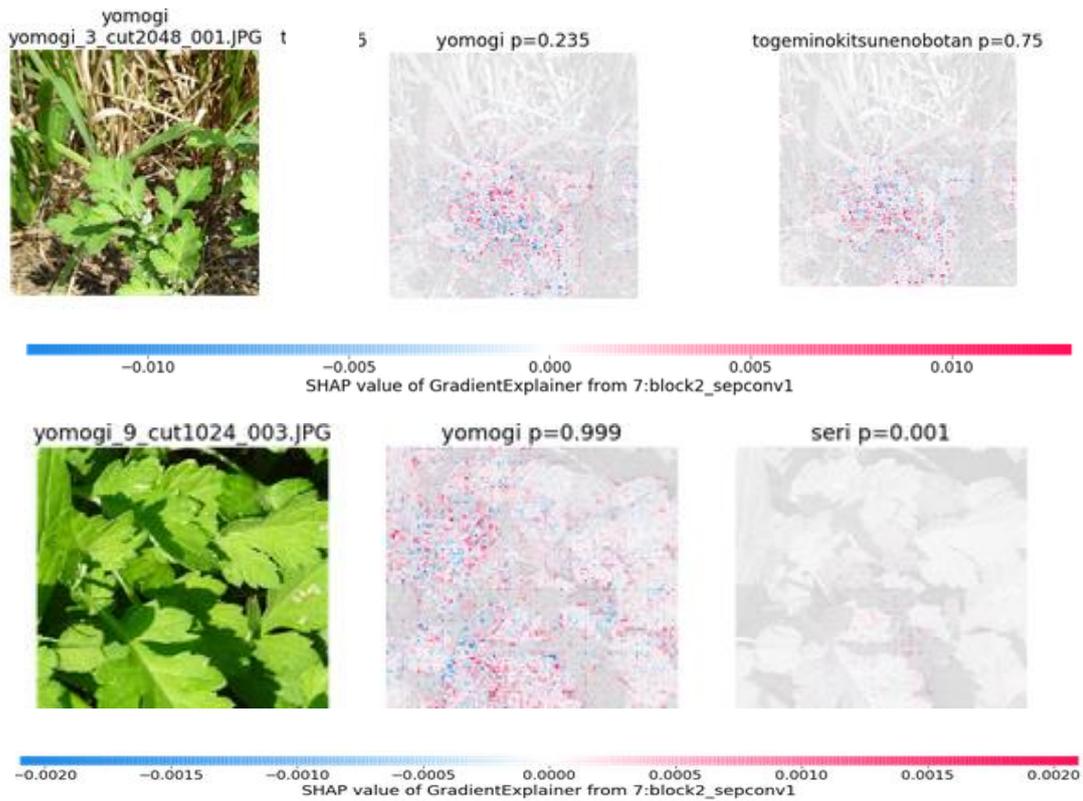


図 4-19 ヨモギ（葉画像）の SHAP 解析結果

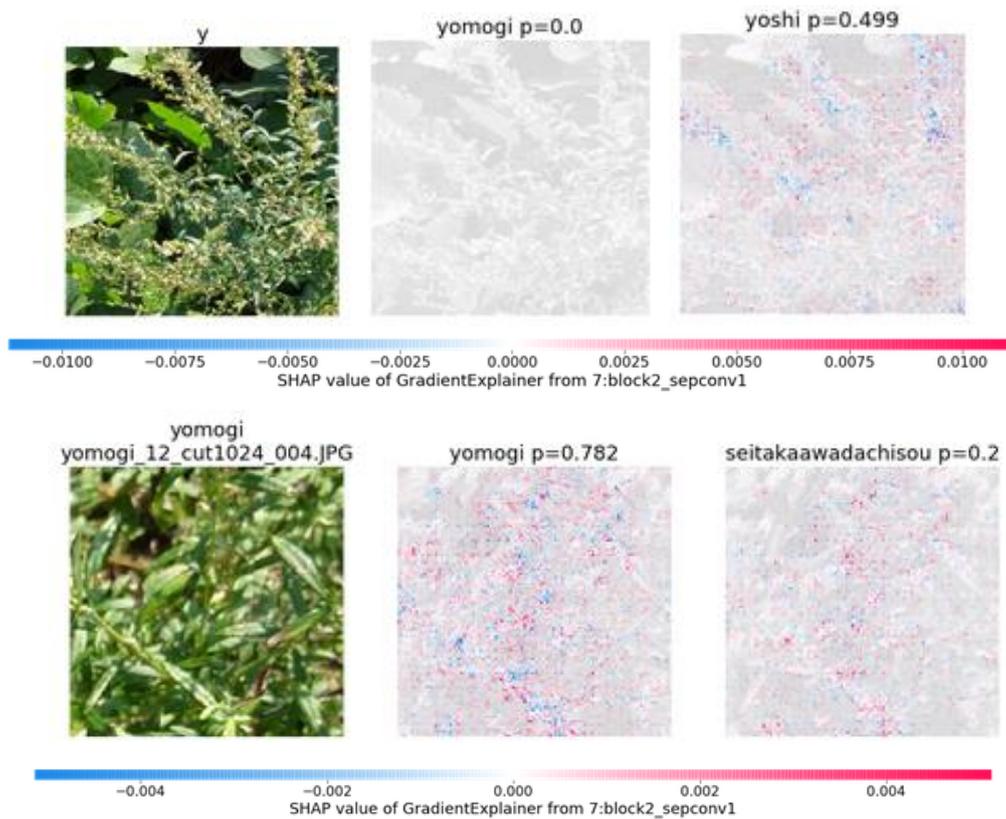


図 4-20 ヨモギ（花画像）の SHAP 解析結果

また、これまでの知見により、学習時の部分画像を作成方法の見直しを行った。まず、画像の端 25% を用いないことで、部分画像にしたときに植物が中心に写りやすくした。そして、部分画像を作成するときに画像サイズの半分以上離して作成することで似たような部分画像が出来にくくした。この方法で作成した部分画像を用いて 2020 春調査の学習モデルを作成しなおすと、別途撮影した 14 種の識別精度は表 4-11 に示すように top3 で 4.7% 向上した。植物種別にみると前モデルで精度が低かったアレチハナガサやヘラオオバコ、ヨモギに加えクソカズラとセイバンモロコシでも確度は向上した。

表 4-11 部分画像作成法改善前後の深層学習結果

	top1	top2	top3
2020 春モデル	65.4	71.0	75.7
新 2020 春モデル	66.4	73.8	80.4

4.3.5 2020 年秋 3 調査地点の深層学習結果

2020 年秋の調査も同様に 3 地点行い、2019 年と 2020 年に行った計 10 回の植物調査で撮影した植物は全 121 種となった。すべての調査画像から構築した画像データセットで学習し識別を行った。さらに別地点で撮影した 14 種との識別も行い表 4-12 に示す。表 4-11 と比較して種数が増えているが確度は向上した。

また、この 14 種の試験では種によって枚数や撮影した季節に偏りがあった。そのため別地点で再度撮影し、植物種を 28 種に増やし、且つ試験画像が 10 枚以上ある種は 10 枚に統一して試験を行った（以後、新別地点画像と呼ぶ）。10 枚に統一する際は、被写体の大きさや明るさ、単体・複数で写っているなど網羅性が出るように選択した。なお、エノコログサとヘラオオバコは撮影時に個体数が少なかったことにより、10 枚未満となっている。表 4-13 に識別結果を示す。ヘラオオバコ、オニノゲシ、クズ、ヒメムカシヨモギ、ヤブガラシ、セイヨウヒキヨモギについては確度が 30% 以下であったため詳細解析を行った。ヘラオオバコは試験画像 3 枚のうち 2 枚が最も高い予測確率でブタナと回答していた。SHAP 解析を行うと花の部分に着目しており、ブタナの蕾と間違えて誤答しているのではないかと考えられる。オニノゲシは、正答した画像が 1 枚も無かったが、科別では 7 枚の画像が top3 でキク科の植物を回答していた。SHAP 解析でも蕾に着目している画像は top3 内にキク科を回答していることから特徴は捉えていると考えられる。クズは 2020 春モデルと比較して確度が下がっている。一方で正解している画像は予測確率 99% 以上でクズと回答しているが、10 枚の画像の中で最も葉が拡大された画像となっていることから生い茂っている画像では葉の特徴を捉えられていないと考えられる。ヒメムカシヨモギでは科別にみても確度

表 4-12 2020 秋モデルの各試験結果

	画像数	正答画像数			確度	
		1st	2nd	3rd	top1	top3
学習時	8398	8255	73	28	98.3	99.5
学習後	4309	4229	47	15	98.1	99.6
別地点 14 種	270	188	16	8	69.6	78.5

表 4-13 新別地点画像を用いた枚数統一試験結果

	画像数	top1	top3
アカツメクサ	10	100.0	100.0
アキノノゲシ	10	100.0	100.0
アメリカウンランモドキ	10	100.0	100.0
アレチハナガサ	10	70.0	100.0
チガヤ	10	100.0	100.0
エノコログサ	7	100.0	100.0
ヘクソカズラ	10	90.0	90.0
ヘラオオバコ	3	0.0	0.0
ヒメジョオン	10	90.0	100.0
ヒメムカシヨモギ	10	0.0	10.0
カラムシ	10	80.0	100.0
カタバミ	10	80.0	80.0
コバンソウ	10	70.0	100.0
クズ	10	30.0	30.0
メマツヨイグサ	10	50.0	80.0
ミヤコグサ	10	90.0	100.0
オニノゲシ	10	0.0	0.0
ランタナ	10	90.0	100.0
セイバンモロコシ	10	80.0	90.0
セイタカアワダチソウ	10	60.0	70.0
セイヨウヒキヨモギ	10	10.0	10.0
スギナ	10	80.0	100.0
タチスズメノヒエ	10	80.0	100.0
タマスダレ	10	100.0	100.0
ツユクサ	10	100.0	100.0
ヤブガラシ	10	10.0	10.0
ヨモギ	10	70.0	80.0
ヨシ	10	80.0	100.0

は上がりず、top3に入った1枚は背景に大きな木があり対象植物に着目しやすくなっていたと考えられる。ヤブガラシはコセンダングサと回答している画像が多かった。これは学習画像を生き茂っているところで撮影したため特徴を捉えられず、学習画像を撮影した場所に似ているコセンダングサと誤答したと考えられる。セイヨウヒキヨモギはグミと誤答している画像が多かった。SHAP解析からこれはセイヨウヒキヨモギだけでなく背景に映り込んでいる木に反応し、グミの枝と間違えているのではないかと考

えられる。画像単位での識別であることから、学習画像、識別画像ともに背景情報による特徴箇所の分散が精度に与える影響は大きいと考えられ、このような画像の識別には物体検出や領域検出が有効と考えられる。

4.4 考察

画像単位の深層学習法では高い精度を得るためには大量の学習用画像が必要で、ILSVRC では 1000 種対象で 45 万枚の画像が用意されている。本章のように独自に撮影した画像を用いる場合には各対象の撮影枚数は少なく、より精度の高い学習法を工夫・改善していくことが重要である。植物種調査への深層学習を応用する上で、本章で得られた知見、指針を以下にまとめる。

- ①撮影した画像は各対象 50 枚程度であっても Keras で用意された神経回路モデル Xception の転移学習で 100 種を超える植物種に対して精度の高い学習が可能である。
- ②撮影した植物の全体画像から細部情報を切り出した部分画像を加えて学習させることが精度向上に有効である。
- ③部分画像は対象が中心に写り、小葉など葉の多い場合は葉が完全に複数写っている画像を選別する。
- ④茎、葉、穂などが細くて背景が写り込む植物は被写界深度を狭くするか、後ろに板を置いて撮影することで背景（他植物や物体）の影響は低減する。
- ⑤SHAP 解析により神経回路モデルが識別時に反応する画像内の箇所が明確になり、精度高低の要因解明に有用である。

本章の結果でも撮影画像が多い種は top1 での精度が高い傾向があり、オニノゲンやヤブガラシは 30 枚程度と撮影画像が少なく精度が低かった。また、セイヨウヒキヨモギは、花画像は多いが撮影回数が少ないため精度は低く、ヨモギは複数回の調査で多様な画像を増やすことで精度は大きく向上できた。撮影画像増強による特徴抽出は有効であり、専門家が判別した質の高い画像の活用など、学習画像入手経路の多様化について検討する。

4.5 本章のまとめ

本章では、①植物種によって精度が大きく低下する要因を解明して植物の撮影や学習方法の改善指針を得ること、②改善指針に基づき新たな植物調査と学習を行い改善の効果を確かめること、③新たな植物調査と学習を行い未知データに対する識別で精度を向上させて汎化性能を高めることを目的とし、全 10 回の植物調査を行った。結果として 121 種の植物画像のデータセットを作成し、98%（上位 3 位では 99.5%以上）の高い精度で分類することが出来た。学習後試験において植物種によって精度が大きく低下した植物は要因を解明することができた。要因解析に基づき、新たな植物調査と学習を行い、汎化性能を高め、新たに精度が低下した植物は要因解析を行うことが出来た。

参考文献

- 1) ImageNet, <http://image-net.org/>, 参照 2022-12-09
- 2) Francois, C. (2017) : Xception: Deep Learning with Depthwise Separable Convolutions, Computer Vision and Pattern Recognition Conference, 2017, pp.1251-1258
- 3) 中山紘喜・曙瑠美・井畔実穂・西野友子・野上敦嗣 (2019) : 深層学習法による高精度植物識別システムの構築と里山保全への適応, 環境システム研究論文発表講演集, 47 巻, pp.43-48
- 4) 白石瑠菜・中山紘喜・西野友子・野上敦嗣 (2019) : 植物調査への深層学習法の応用に関する研究, 環境システム研究論文発表講演集, 47 巻, pp.103-107
- 5) 響灘ビオトープ, <http://www.hibikinadabiotope.com/>, 参照 2022-12-09
- 6) 西野友子 (2019): 北九州市内河川公園の植物相とその特徴, わたしたちの自然史特別号, 初版, pp.31-62
- 7) 竹内均 (2002): 地球環境調査計測事典第 1 巻 陸域編②, 株式会社フジ・テクノシステム, 初版, p.263
- 8) 竹内均 (2002): 地球環境調査計測事典第 1 巻 陸域編①, 株式会社フジ・テクノシステム, 初版, pp.713-714
- 9) 国土交通省 : 国土技術政策総合研究所, http://www.nilim.go.jp/lab/fbg/ksnkankyō/mizukokudam/system/download/H28D_Chousamanual_dam/H28D_07shokubutu.pdf, 参照 2022-12-09
- 10) Keras : <https://keras.io/>, 参照 2023-02-03
- 11) Erik, Š. And Igor K. (2014) : Explaining Prediction Models and Individual Predictions with Feature Contributions, Knowledge and Information Systems, vol.41, pp.647-665
- 12) 林弥栄 (2002) : 野に咲く花, 株式会社山と溪谷社, p.288

第5章 深層学習法による植生分布の解析

5.1 はじめに

前章では深層学習法の植物識別への応用について検証した。本章では前章で得た知見を基に、深層学習法の植物保全活動への応用として、木本種および草本種の分布解析について述べる。草本種の保全管理における課題は前章で触れているが、木本種すなわち森林環境の保全についても類似した課題がある。従来人によって整備されてきた里山は、森林環境の保全は生物多様性や環境教育等の観点から重要である。日本国内では森林生態系多様性基礎調査が5年をかけて平成11年度より4回にわたり行われている他¹⁾、地方各地の放置された里山に対し、地方自治体や市民団体が保全活動を行っている²⁾。これらの調査や保全活動において、樹種組成の把握は重要な要素だが、そのためには専門家による現地踏査のように多くの人手と労力がかかり、調査頻度や地域の制限から十分な調査が行われていない。このことから調査、識別の効率化を行う手法について研究することは草本種同様大きな意義がある。また、近年では小型無人航空機（UAV）を用いて、広域かつ短時間で画像データを収集することができるようになった。これは木本種、草本種どちらの調査においても有用である。本章では木本種、草本種を対象に、構築した識別システムの植物保全管理への適用性の検証を目的として、以下の3つに取り組んだ。①深層学習法を用いた樹冠識別から植生分布を求める手法の構築、②里山樹木に対して、高い確度を得る画像群の撮影条件の検討、および本システムを用いて作成した樹木分布図の里山管理への適用性の検討、③草本種を対象に前章において構築した学習モデルを用いて俯瞰画像から植物分布図の作成、および植生管理への適用性の検討。

5.2 植生分布の解析手法

5.2.1 調査地区

木本種の調査は専門家同伴の下、北九州市立大学ひびきのキャンパス構内の自然保護地区にある里山を対象に行った。この地区は1995年から始まった北九州学術研究都市整備第1期事業の中で、2000年に整備された。元々は全て切り崩す計画だったが途中で変更し、約21000m²の地区を人の手を加えない形で残したもので、人による管理はされていない。里山の地上からの高さは約40mである。樹木は見る角度や距離によって撮影される姿形が大きく異なり、葉に限定しただけでも下から見上げる、真上から見下ろす、斜め上から俯瞰して見る、近くから葉の形が分かるように見る、遠くから葉の塊を見るなど様々なケースが想定される。1つの学習モデルでこれら多種の姿形を持つ樹木を識別するためには、学習データに多様な角度、距離から撮影した写真を含める必要がある。そこで樹木撮影では、デジタルカメラを用い下方から正面を捉えた画像と、ドローンを用い上方から見下ろした画像を撮影した。撮影機材はSONY製デジタルカメラDSC-HX60Vと、DJI製ドローンMAVIC2 Proを用いた。撮影した画像の解像度はそれぞれ、5184×3888、5472×3648である。さらにドローンでは里山全域に対して、真下を見下ろす角度で樹冠を収めたものを撮影した。また、草本種における俯瞰画像の撮影条件を表5-1に示す。

表 5-1 俯瞰画像の撮影条件

	俯瞰画像 1	俯瞰画像 2
撮影日	2020/5/19	2021/3/8
撮影場所	撥川河川敷	響灘ビオトープ
使用機器	デジタルカメラ Panasonic DC-FZ85	ドローン DJI MAVIC2Pro
画素数	4896×3672	5472×3648
撮影高度	1.7m	3m

5.2.2 深層学習方法

解析に用いる学習モデルの構築には、深層学習フレームワークにKeras³⁾/Tensorflow、神経回路網モデルにXception⁴⁾を用いた。深層学習時のハイパーパラメータは、エポック数を20、バッチサイズを16とし、学習率は0.01から5エポック毎に減衰率0.6で小さくするよう設定した。学習はカラー画像で行い、Kerasに搭載されているデータ増強処理を用いた。学習画像の振り分けは、木本種が7.5:1.5:1の割合で学習、検証、試験へそれぞれランダムに振り分けた。さらに、ExplainableAIの1種であるSHAP⁵⁾を用いて識別時における画像の注目領域を解析した。草本種に関しては第4章で構築した学習モデルを流用している。目標識別精度については第4章同様に、表3-1より学習に用いたXceptionのTop1確度とTop5確度の平均値86.7%を種ごとの識別においてTop3確度で目指す。

5.2.3 学習画像と試験画像の作成

撮影した画像に画像処理を施し学習画像を作成した。Xceptionを用いる場合、入力画像サイズは299×299となるので、入力画像のアスペクト比を1:1に揃えることで、学習時の画像の歪みを防いだ。撮影画像の縦横のうち短い方にサイズを揃え、中心から切り出すことでアスペクト比を揃えた。また撮影画像から1024×1024と2048×2048の部分画像を切り出し学習画像に加えることで、樹木の枝葉情報を学習できるようにした。

ドローンで撮影した画像のうち、樹冠画像について学習画像には加えず、個々の樹冠を切り出し試験画像とした。樹冠の切り出しについては、webサービスDroneDeploy⁶⁾を用いて19968×12544のパノラマ画像を作成した後、以下の方法で抽出した。まずMeanShift法⁷⁾を用いて、画像をスーパーピクセル化し、樹冠ごとに区分した。次に、樹冠ごとにスーパーピクセルで区分した画像に対し、Slic法⁸⁾を用いて区分された領域を1つずつ切り出した。さらに、切り出した樹冠に外接及び内接する矩形を抽出し、樹冠データとした。領域分割についてはSlic法のみでも可能だが、事前にMeanShift法を用いることで樹冠境界が鮮明になり、適切な範囲の領域分割が可能となる。1つのポリゴンに複数の樹冠が含まれる場合は再度上述した分割法を用い、1つのポリゴンに含まれる樹冠がなるべく1種になるようにした。これらの加工は、JupyterNotebook上で画像処理ライブラリのOpenCVおよびscikit-imageを用いた、自製pythonプログラムにより行った。樹冠を囲む外接矩形は、OpenCVに用意されているboundingRect関数で簡単に求めることができるが、内接する矩形を求める関数は画像処理ライブラリには用意されていないため、次の手順

のプログラムを作成した。①樹冠の輪郭線からポリゴンの重心を求める。②重心から上下左右の4区画に輪郭線を分割し、重心から各区画の最小値を求める。③上下左右の最小値を長方形のYmin, Ymax, Xmin, Xmaxとする。

草本種においてはドローンを用いて、撥川で撮影した斜め下を俯瞰した画像と、響灘ビオトープで撮影した真上からの画像を対象に、MeanShift法で400×400～500×500程度のポリゴンに領域分けし、各領域の内接矩形画像に対して学習済みモデルによる植物識別を行い、同種の隣接するポリゴンを結合することで、植物分布図を得た。ポリゴン分割の例を図5-1に示す。



図5-1 ポリゴン分割及び識別画像の抜き出し

5.3 調査、学習結果および分布図の作成

5.3.1 木本種用画像データベースと試験画像の作成

学習対象の樹木種は里山を踏査した際に確認できた樹木のうち、広域に分布しているもの、葉の色が似ているもの、特徴的な形状をしているもの、撮影が容易だったものを中心に選び、合計で8種の樹木を撮影した。ドローンの撮影は高度約50メートルから行った。撮影画像に画像処理を加え学習画像データベースを作成した。その内訳を表5-2に示す。アカメガシワは他の樹木と比べ樹高が低く、ドローン撮影では他の樹木に覆われる範囲が多かったため、ドローンによる撮影は行わなかった。

樹冠の撮影は、高度80メートルから行った。これを基にパノラマ写真を作成し、樹冠画像の切り出しを行った。その様子について図5-2、図5-3に示す。切り出した樹冠画像の中から識別に用いる8種のうちアカメガシワを除いた7種について試験に用いる樹冠画像を選び出した。

表5-2 撮影画像と画像データベースの内訳

種名	撮影数		部分画像数		学習画像数
	デジカメ	ドローン	1024	2048	
アカマツ	104	38	1208	800	2150
アカメガシワ	99	0	872	583	1154
アラカシ	110	2	1068	654	1834
ハチク	20	13	188	304	525
ヒメユズリハ	102	104	1940	1207	3353
クロキ	28	6	243	151	428
ツブラジイ	104	11	961	753	1829
ヤマモモ	43	23	631	336	1033



図5-2 里山全域を収めたパノラマ写真



図5-3 切り出した樹冠画像の一例

5.3.2 学習と試験における確度

作成した学習モデルによる学習時の確度と、樹冠画像を用いた試験時の正答数を表5-3、表5-4に示す。top1は予測確率1位のみの正答数、top3は予測確率3位までに正解が含まれている場合の正答数を示す。表5-3より、学習時の確度は全体を通して90%以上の高い確度が得られたが、表5-4より、樹冠識別についてはアカマツのみがTop1で90%弱、他の樹木は正答数が0だった。Top3まで含めた平均確度は50%弱で、Top3まで含めたとしても、全体的に確度は低かった。学習時の確度は高く、樹冠識別時の確度が低かった原因としては、学習時のオリジナル画像が全体的にデジタルカメラのものが多く、樹冠のような上方から見た画像の学習が不十分だったことが考えられる。その上で、アカマツの確度が樹冠識別でも高かった要因としては、他種と比べ葉が特徴的な形状をしており、学習時のデータ数も多く、質の高い学習データが用意できたことが影響していると考えられる。

表5-3 里山樹木の種ごとの学習時における確度

種名	確度
アカマツ	100.0%
アカメガシワ	99.4%
アラカシ	96.8%
ハチク	100.0%
ヒメユズリハ	99.2%
クロキ	100.0%
ツブラジイ	100%
ヤマモモ	92.3%

表5-4 樹冠画像識別テストの結果

種名	試験数	top1	top3
アカマツ	9	8	9
アラカシ	5	0	1
ハチク	10	0	10
ヒメユズリハ	8	0	1
クロキ	10	0	6
ツブラジイ	8	0	1
ヤマモモ	14	0	5

この結果を受け、すべての樹冠画像に対し予測確率が1番高い種を調べたところ、約78%をアカマツと識別していた(表5-5)。調査対象林において、アカマツは図5-2の右側一部分に育成しており、この結果から、全体的に樹冠画像をアカマツと誤認していることが分かった。アカマツは図5-4のように針状の葉が密集しており、入力時の画像は299×299まで縮小されることから、画像の大部分を輪郭の不鮮明な緑が占め、特徴の潰れた樹冠画像をアカマツと誤認しやすい状況になっていたと考えられる。

表5-5 樹冠画像識別テストにおけるTop1回答数

種名	Top1 回答数
アカマツ	402
アカメガシワ	1
アラカシ	13
ハチク	24
ヒメユズリハ	35
クロキ	33
ツブラジイ	7
ヤマモモ	2



図5-4 アカマツの学習画像の一例

そこで誤答した樹冠画像についてSHAP解析を行い、注目領域を調べた。その一例を図5-5に示す。注目領域は図5-5のように樹冠全体に点在する傾向があり、樹木由来の特徴を判断しているのかSHAPでは判明しなかった。しかし樹冠画像は外接矩形を用いているため、図5-5のように注目領域が樹冠部分に集中することで、背景部分から樹冠の形状が浮き上がる傾向を確認できた。これは樹冠の形状が識別に影響していると考えられ、その影響を除外するため試験画像を外接矩形から内接矩形に変更した(図5-6)。また樹冠画像のアスペクト比が1:1でなかったため1:1に整えた。



図5-5 SHAPによる樹冠の判断箇所解析

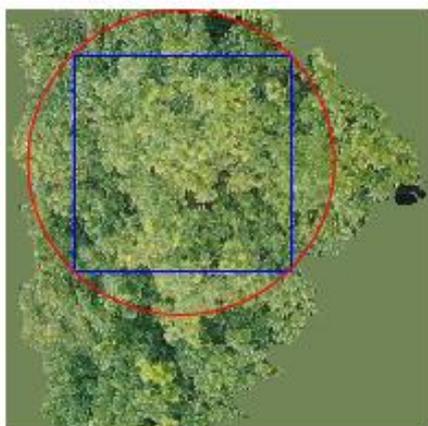


図5-6 内接樹冠画像の切り出し

内接矩形による樹冠識別の結果を表 5-6 に示す。表 5-4 と表 5-6 より top3 においては、ハチクのように識別できていない種もあるが、ヒメユズリハ、クロキ、ツブラジイにおいては正答数が増えており、top1 においても、この 3 種は正答数が増加している。このことから内接矩形を用いて樹冠の形状を除外することは精度向上に繋がると言える。しかし top1 の正答数は低い傾向にあり、精度向上には画像データベースの改善が必要なものと考えられる。

表5-6 内接樹冠画像識別テストの結果

種名	試験数	Top1	Top3
アカマツ	9	7	9
アラカシ	5	0	1
ハチク	10	0	0
ヒメユズリハ	8	1	5
クロキ	10	3	10
ツブラジイ	8	4	8
ヤマモモ	14	0	5

5.3.3 画像データベースの再構築と確度検証

画像データベースとの問題点としては、学習画像のうち部分画像について、同じような範囲を切り出した画像を使っていることが考えられる。また試験に用いた樹冠画像の解像度が低く細部が潰れており、この改善も精度向上へ繋がると考えられる。これらの問題点を解決するため新たに画像データベースと樹冠画像を作成した。樹冠画像については、高度50メートルで撮影した画像から樹冠の内接矩形を切り出した。パノラマから切り出した樹冠と、撮影画像から切り出した樹冠の比較を図5-7に示す。図5-7より解像度の向上については違いがはっきりでない結果となった。これについては撮影時ドローンを移動させながら撮影したことや、風によって樹木が揺らいだことによるブレによるもので、今後データを収集する際に注意すべき点と言える。

画像データベースについては、画像を整理した上で、部分画像の追加を段階的に行い、計7個の画像デ

データベースを用いて確度の改善を行った。作成した画像データベースの内訳を表5-7に、top3における確度の変化を図5-8に示す。1回目の改善では、デジタルカメラで撮影した画像について、部分画像のすべてを削除した。本研究では樹冠を識別しているのので、枝葉を下方から撮影した画像を増やすことは不適切であると判断した。しかし多様な角度からの識別を考慮し、部分画像の元となったオリジナル画像は学



図5-7 新旧樹冠画像の比較

表5-7 画像データベースの内訳の変化

改善番号	アカマツ	アカメガシワ	アラカシ	ハチク	ヒメユズリハ	クロキ	ツブラジイ	ヤマモモ
1	330	99	136	190	693	81	248	174
2	329	99	143	322	1144	120	357	274
3	329	99	143	323	497	120	357	275
4	479	99	173	518	752	180	522	425
5	479	99	173	518	1144	180	522	425
6	380	99	153	387	581	141	412	324
7	480	99	173	517	751	181	522	424

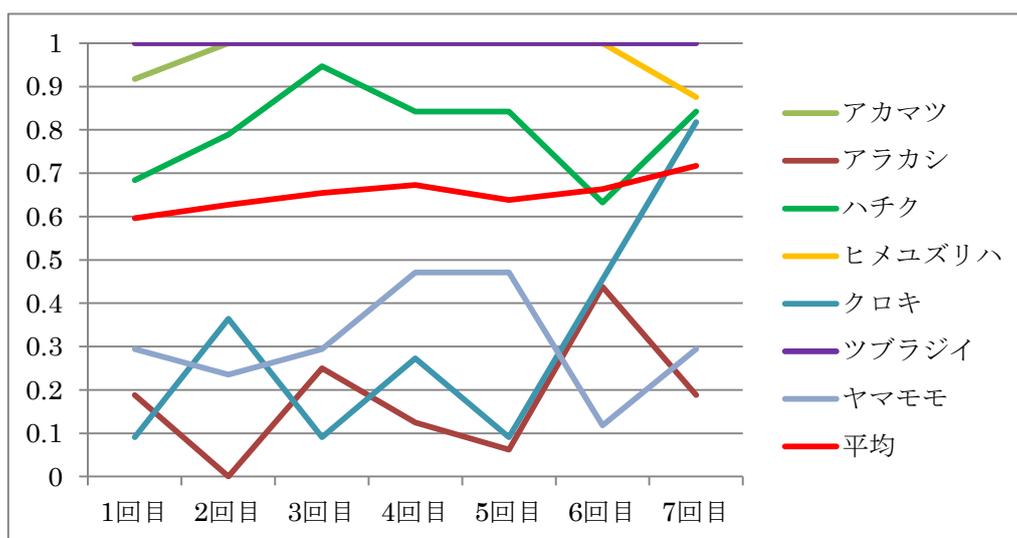


図5-8 画像データベース改良によるtop3確度の推移

習画像に残した。ドローンで撮影した画像については、同じ樹木に対して似た角度から撮影した画像を削除した上で、サイズが512×512, 1024×1024, 2048×2048となる3種類の部分画像を、オリジナル画像数に対しそれぞれ、アカマツが10枚, 10枚, 2枚, 残りの樹木は5枚, 5枚, 2枚で切り出した上で整理した。top1の正答数がアカマツ以外低かったので、アカマツを除く樹木の画像数の変化量を大きくした。その内訳を以下に示す。2回目は1回目に対し、アカマツを除く樹木の部分画像のうち、512×512と1024×1024の切り出し数をアカマツ同様10枚にした。3回目は2回目に対し、画像数の多いヒメユズリハのオリジナル画像を半分にした。4回目は3回目に対し、新たに学習時の入力サイズと同じサイズである、299×299の部分画像をオリジナル画像1枚につき15枚切り出して加えた。5回目は4回目に対し、ヒメユズリハのオリジナル画像の枚数を、2回目までと同じ枚数に戻した。6回目は5回目に対し、ヒメユズリハのオリジナル画像の枚数を再び半分にし、299×299の部分画像を全て除去し、代わりに512×512の部分画像の切り出し数を15枚に増やした。7回目は6回目に対し299×299の部分画像を10枚ずつ切り出した物を加えた。図5-8より7回目の改善における平均の確度が一番高くなった他、2～3回目の確度の変化から、ある種類の学習画像数を変化させると他種の確度にも影響が出ていることが分かった。しかし、アラカシとヤマモモについては確度が低いままで、画像データベースの更なる改善が必要と思われる。

この結果を受け、試験に用いた樹冠画像とは別の樹冠画像を各種10枚ずつ画像データベースに加えて学習させ、樹冠識別を行った。その結果を表5-8に示す。この結果、アラカシを除く6種で90%を越す確度を達成できた。これは、試験画像と同程度の距離と角度で撮影された画像を学習に用いることで、確度が向上したものと考えられる。

表5-8 ドローン内接樹冠画像による試験

種名	試験数	Top1	Top3
アカマツ	12	12	12
アラカシ	16	8	11
ハチク	19	15	19
ヒメユズリハ	8	8	8
クロキ	11	5	11
ツブラジイ	16	13	15
ヤマモモ	17	12	17

5.3.4 樹木分布図の作成

前節までで得られた知見から、この画像データベースより作成した学習モデルの活用法の1つとして、里山の樹木分布図を作成した(図 5-9, 図 5-10)。図 5-2 から分割したすべての樹冠に対し、構築したシステムを用いて種の識別を行い、隣接樹冠が同種ならば1つにまとめた。里山の樹木は本研究で用いた8種以外にも多種類存在するため不完全なものであり、樹冠識別に用いた7種についても一部誤答が確認できた。これは識別できる種数に起因すると考えられ、今のままでは適用範囲が限定されている。適用範囲の拡大には識別種の追加が必要である。また Xception の入力サイズが 299×299であることを考慮した上で、樹冠画像の高解像度化についても検証の必要がある。

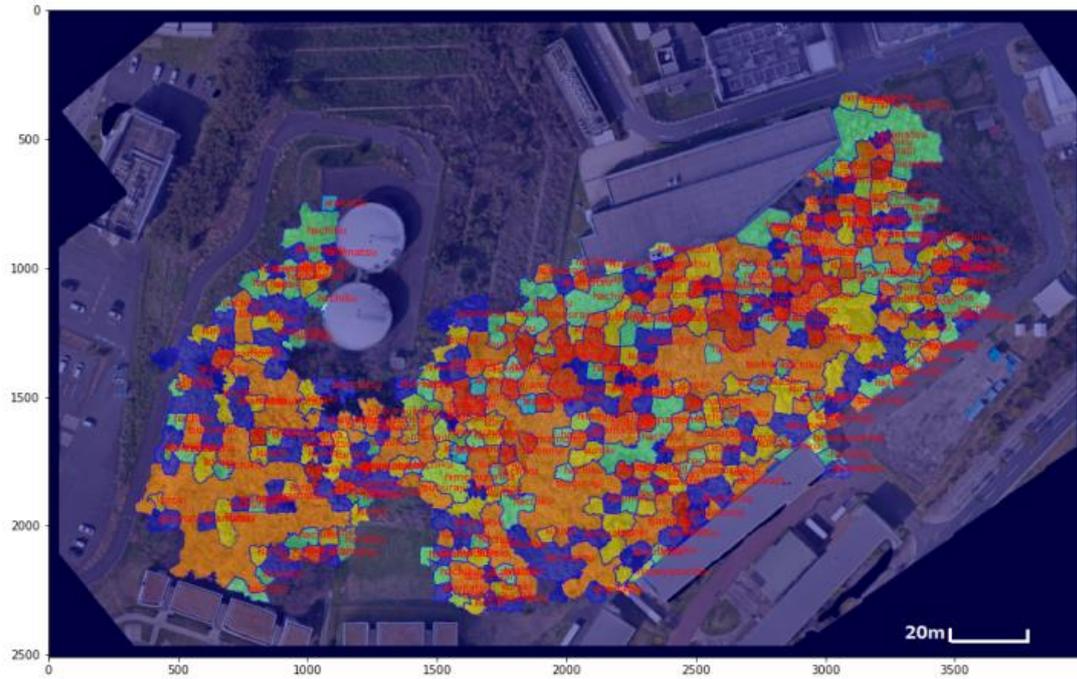


図5-9 里山樹木分布図尺度20m

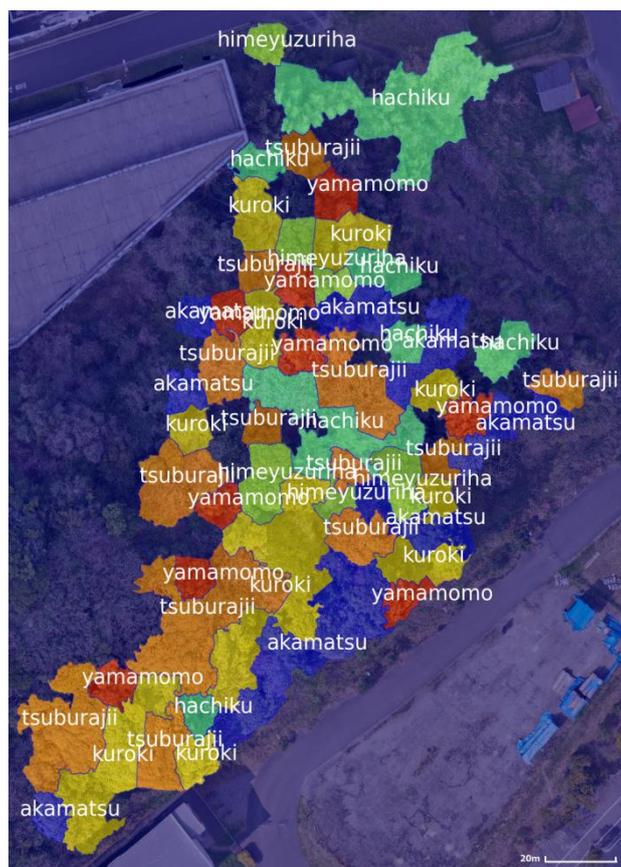


図5-10 里山の樹木分布図尺度20m(一部拡大)

5.3.5 俯瞰画像からの草本種識別結果

前節の結果より草本種においても同様の分布解析を検証した。第4章で最終的に構築したモデルを用いて俯瞰画像1とそれを最大50分割に領域分割したものを図5-11, 植物種別に領域分けしたものを図5-12に示す。なお、正誤の判断は撮影時に同伴した専門家がを行い、正答した種には○, 誤答した種には×, 画像からは正答しているか判断が出来なかった種には△を記載した。人間の目で見てもわかるようなアカツメクサやシロツメクサ, ヨシは正しく識別し領域分け出来ていることから, 俯瞰画像からの種の識別も可能であると検証できた。一方で手前のアカツメクサは正しく識別できていたが, 奥のアカツメクサはアレチハナガサだと誤答した。これはデジタルカメラで撮影したために斜めから撮影した画像になっているため同じ種でも大きさが違うように見えてしまっていることが起因していると考えられる。

この結果を受け, ドローンを用いて真上から撮影した俯瞰画像2とそれを最大30分割に領域分割したものを図5-13, 俯瞰画像1と同様の学習済みモデルを用いて分類を行った結果を図5-14に示す。真上から撮影した俯瞰画像では, ポリゴンの分割数を最大50分割にした場合, ポリゴン面積が小さくなり, ポリゴンの内接円から取り出した識別画像に写る植物が拡大されていたため, 正しい識別は出来なかった。そこで分割数を最大30分割にすることで, 1ポリゴンの大きさを800×800~1000×1000にして識別を行うと, 図5-14に示すようにヨシとアカツメクサの群落を識別するようになった。ドローン画像による俯瞰画像の識別は, 学習画像に真上から撮影した画像が少なかったことより俯瞰画像1と比較して識別されにくかったと考えられる。



図 5-11 俯瞰画像 1 とその領域分割

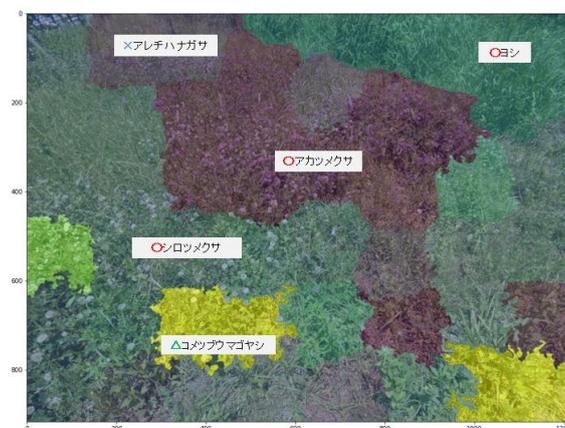


図 5-12 俯瞰画像 1 の Xception モデルによる領域分け結果
(正答=○, 誤答=×, 判断不可=△)

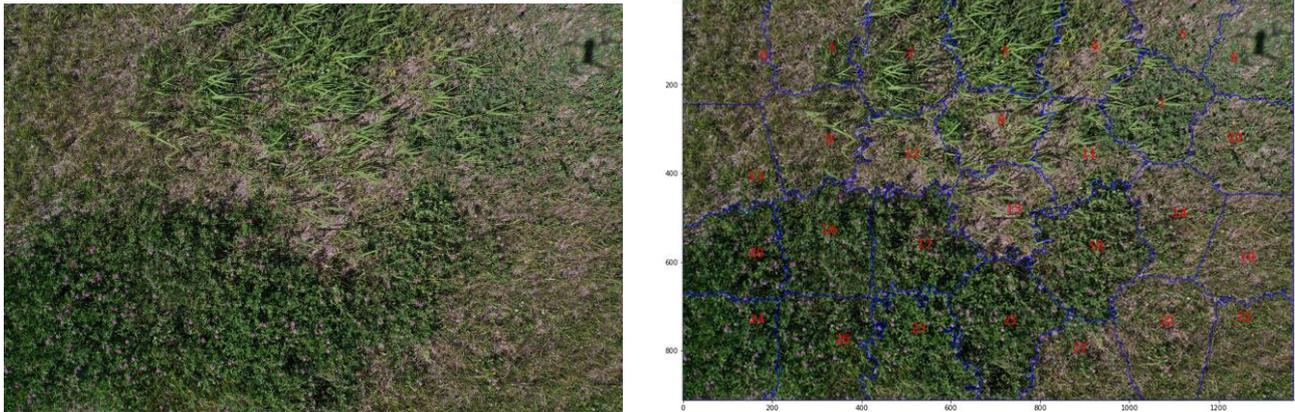


図 5-13 俯瞰画像 2 とその領域分割

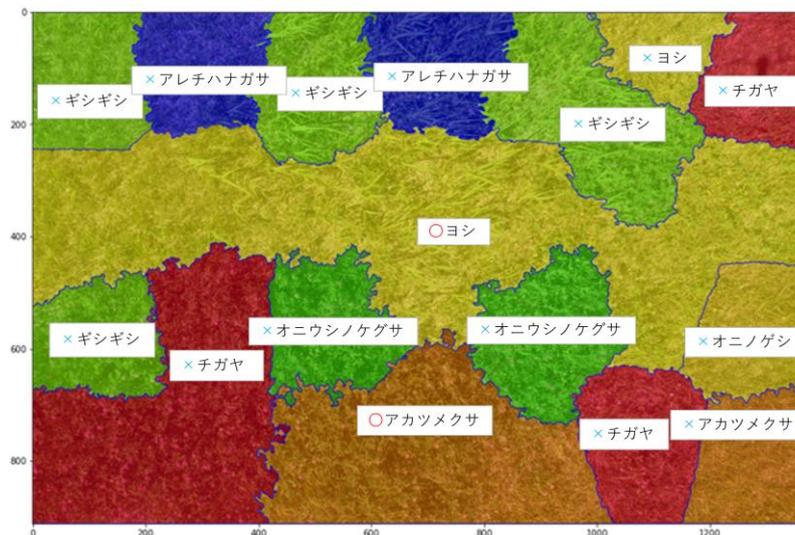


図 5-14 俯瞰画像 2 の領域分け結果

5.4 考察

表5-2, 表5-4と図5-8を比較すると, 画像数は改善した画像データベースの方が少ないが, 確度は改善後の方が高くなっている. これは学習モデル構築には学習枚数と同程度に画像の質も重要であることを示している. 学習画像から似た構図の画像を排除し, 画像のアスペクト比を学習時の入力画像と同じ1:1に保つことで確度の向上につながると考えられる. しかし, 学習画像に対しこれらの改善を行っても本研究ではTop3確度の平均は70%だった. これに対し学習画像に樹冠画像を混ぜて学習させた結果, 表5-8のようにTop3確度の平均は90%を超えた. 本研究では様々な角度からの識別が行えるよう多角的な学習画像を用いたが, 樹冠識別においては, 斜め上からの画像だけでは十分な学習が行えず, 樹冠画像自体の学習が必要であることが分かった. 最終的に多角的な撮影から構築した学習画像データベースを用いて, 高い確度を持つ樹冠識別を行うことができたが, 樹種による画像枚数や, 撮影角度の比率にばらつきがある. 表5-7と図5-9より, ある種の画像枚数の増減が, 別の種の確度に影響を与えていることから, これらが全体の確度に与える影響について, 今後検証する必要がある. また, 樹冠を収める距離での撮影は天候や時間帯により葉の階調が異なる. 本研究で作成した樹冠画像は, 樹冠の形状に内接する矩形を切

り出すことで形状の影響を極力取り除いたが、階調の違いは考慮できていない。このことから、本研究で作成した学習モデルを用いて、時間帯や場所の異なる樹冠識別を行った場合に高い確度を得ることはできないと思われる。これについては撮影時間、撮影場所、データ増強時の階調の変更による確度の影響を検証する必要がある。

UAVと機械学習による樹木識別については既に中根らによる、グリッド画像とDL法を用いた研究⁹⁾や、鈴木らによる、スーパーピクセルとSVMを用いた研究¹⁰⁾などが報告されている。中根らの研究では、100×100ピクセルのグリッドで画像を取り出すため、2種以上の樹木や背景などが含まれる画像が一定の割合で存在する。そのため竹、ヒノキ、広葉樹の3種に分類に対して、竹、ヒノキ、広葉樹、竹-ヒノキ混合、竹-広葉樹混合、ヒノキ-広葉樹混合、影と7種類のクラスを作っており、樹冠単位のクラス管理が難しい。樹木種が増えた場合は混合クラスの数が増大するため適用が困難と思われる。鈴木らの研究では、スーパーピクセル分割は本研究と同じで、画像の特徴量としてスーパーピクセル中のRGBのヒストグラムを学習させている。したがって隣接する植物種の混合はないが、色情報の分布が近い植物種では識別が困難となる。論文ではアブラガヤ、ハンノキ、イ、ミズソバ、ススキ、ヨモギ、水面の7クラスに分類しているが、色情報のヒストグラムだけで植物種を大きく増やすことは困難と思われる。本手法では、樹冠単位でその内接円から取り出した画像を学習させるため、他の樹木や樹木以外の背景の混合は発生せず、樹木単位のクラス管理が容易である。また、画像を学習させているので、RGBヒストグラムのみでの学習と比べ、樹種の拡張も容易と思われる。

最後に木本種、草本種においてそれぞれ構築した学習モデルを用いて分布図を作成した。木本種については調査区の総樹種に対し調査樹種は少数にとどまっており、多くの樹種が識別できる学習モデルを構築することで改善が期待できる。識別できる樹種が増えることで、樹種自体の識別による分布図ではなく、針葉樹や広葉樹、竹といった大まかな区分による識別においても精度の高い分布図の作成が可能と考えられる。これは草本種においても同様で、植物調査において人手による現地踏査の前にドローン等を用いた事前調査を行い、大まかな分布図が作成できれば、現地踏査の負担を軽減する一助となることが期待される。この場合分布図には位置情報を含める必要があるが、分布図のオルソ化などにより対応可能と思われる。また領域分割の精度改善も課題であり、深層学習法による領域検出手法であるセマンティックセグメンテーションの適用などの検証が必要である。

5.5 本章のまとめ

本章では木本種を対象とし、①深層学習法を用いた樹冠識別から植生分布を求める手法の構築、②里山樹木に対して、高い確度を得る画像群の撮影条件の検討および本システムの里山管理への適用について、樹木分布図の作成を目的とし里山樹木の調査を行った。その結果デジタルカメラとドローンを用いて、多様な角度で撮影したオリジナル画像から樹木8種を高精度に識別する学習モデルを作成できた。8種のうち樹冠を撮影できた7種について樹冠識別を行うと、top3の確度は約50%だった。アスペクト比の統一や類似した部分画像の排除に加え、試験画像と同程度の距離と角度で撮影した画像を含めた学習を行うことで、7種のうち6種の確度が約90%に向上した。構築した学習モデルを用いて樹木分布図を作成した。草本種にもいっても同様に分布図を作成し、双方において領域分割、識別ともに一定の成果は得られたが不十分な点も見受けられた。この改善にはよりよい学習モデルの構築が必要であり、学習画像の枚数や撮影角度の比率、時間や場所の変化が確度に与える影響、領域検出手法の適用を検証する必要

がある.

参考文献

- 1) 林野庁 : <https://www.rinya.maff.go.jp/j/keikaku/tayouseichousa/>, 参照 2023-02-03
- 2) 伊東宏樹・大住克博・奥敬一・衣浦晴生・黒田慶子・高畑義啓・松本和馬 : 里山に入る前に考えること, http://www2.kobe-u.ac.jp/~kurodak/satoyama2010_3.pdf, 参照2023-02-03
- 3) Keras : <https://keras.io/>, 参照 2023-02-03
- 4) Francois, C. (2017) : Xception: Deep Learning with Depthwise Separable Convolutions, Computer Vision and Pattern Recognition Conference, 2017, pp.1251-1258
- 5) Erik, Š. and Igor K. (2014) : Explaining Prediction Models and Individual Predictions with Feature Contributions, Knowledge and Information Systems, vol.41, pp.647-665
- 6) DroneDeploy : <https://www.droneDeploy.com/>, 参照 2023-02-03
- 7) D. Comanicu and P. Meer, (2002) : Mean Shift: a Robust Approach Toward Feature Space Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, No.5, pp.603-619
- 8) Radhakrishna, A., Appu, S., Kevin, S., Aurelien, L., Pascal, F. and Sabine, S. (2012) : SLIC Superpixels Compared to State-of-the-art Superpixel Methods, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.34, No.11, pp.2274-2282
- 9) 中根英昭・若槻祐貴 (2018) : 環境分野への深層学習応用研究の立ち上げについて, 高知工科大学紀要, 15(1), pp. 111-120
- 10) 鈴木太郎・土屋武司・鈴木真二・山場淳史 (2016) : 小型 UAV による空撮画像からのスーパーピクセル分割と機械学習を用いた植生分類手法の構築, Journal of The Remote Sensing Society of Japan, vol.36, pp.59-71

第6章 深層学習法による鳥類のリアルタイム動画解析

6.1 はじめに

前章までで、植物調査および保全に対する深層学習法を用いた画像識別の応用について検証した。本章では、鳥類調査に対する深層学習法の応用について検証する。福岡県北九州市に位置する響灘埋立地は、1980年代より廃棄物処分に伴う埋め立てが始まり、当初は産業用地としての計画がなされていた。しかし廃棄物の埋め立て後に出来た凹凸の地形が、湿地や淡水池、草原などの多様な環境を生み、そこにさまざまな生物が生息するようになった（響灘ビオトープ HP より）¹⁾。その中には、絶滅危惧種に分類されるチュウヒ（タカ科の鳥類、国・絶滅危惧 I B 類、国内希少野生動植物）やベッコウトンボ（昆虫類、国・絶滅危惧 I A 類、国内希少野生動植物）²⁾の生息や繁殖が確認され、希少種を含めた生態系が確立されていることが分かった。北九州市はこれらを保護するため、用地の一部を響灘ビオトープとして整備し2012年に開園した¹⁾。中でもチュウヒは、我が国では主に北海道や東北地方で繁殖しているが³⁾、響灘埋立地においても繁殖が確認されている⁴⁾。チュウヒの営巣・採餌の場所となる水辺のヨシ原・湿地・草地の面積減少や未管理により、繁殖が困難になりつつある状況において、湿地生態系の頂点にいるチュウヒの行動や生息環境を定量的に把握することは、今後本地区の生態系保全対策において極めて重要である。

その一方で本地区での開発は、リサイクル促進工場や再生可能エネルギーをはじめとした発電施設の建設が進み、数年後には洋上風力発電施設の建設も計画されている。このような状況にあって、開発事業に伴う環境アセスメントの実施はより重要となり、過去10数年にわたり13件実施された⁵⁾。このように響灘ビオトープはチュウヒを含む湿地生態系に加え、風力発電施設を含む開発区が隣接した自然保護区という性質を持ち、継続的なモニタリングによる鳥類の生活行動の把握はより重要となる。しかし自然環境を管理する手法のマニュアル化は難しく、継続的なモニタリングにも膨大なコストがかかるため、本地区での実施は見送られているのが現状である。従来鳥類調査は人手によるラインセンサス法、スポットセンサス法、録音、捕獲^{6,7)}などの他、レーダーを用いた調査⁸⁾などがあるが、その多くはコストや人材の面から実施場所や頻度に制限がかかる。これらの課題に対し、野生動物のカメラ観察は定常性において有用だが、膨大な映像データを人手で詳細に解析することは困難であり、近年では機械学習、特に深層学習を応用した画像解析の自動化研究が活発である。中でもYOLO(You Only Look Once)⁹⁾は高速性と高精度を両立しており、映像のリアルタイム解析が期待できる。

響灘ビオトープではチュウヒなど猛禽類の採餌活動が確認され、これらを常時観察するために4Kの全方位マルチセンサーネットワークカメラが1台設置された。鳥類活動の調査には映像の飛跡解析が重要であるが、本章の予備的研究として画像処理法による動体追跡について検証を行った¹⁰⁾。結果として、背景差分法による動体の自動検出と、BOOSTING法¹¹⁾による動体の自動追跡を組み合わせることで、検出鳥類を正確に追跡できることを確認した。しかし、独立した手法の組み合わせであるため、一度追跡が途切れた鳥類の再追跡が困難であること、背景差分法では映像内の動体の中から鳥類のみを抽出することが困難であるなどの課題が見つかった。深層学習法の適用により、検出物を人の判断によらずシステム内で識別することで、検出動体から鳥類の自動抽出が可能になると考えられる。また、ステレオ撮影された2つのビオトープ映像から、カメラとの位置関係が判明している風車までの距離を映像間の視差より求めた。これにより基線長18m、焦点距離32mmの条件下で実測値945mに対し1082m、実測値787mに対し937mの結果が得られ、約15%~20%の誤差が生じていることが分かった。これは飛跡解析

への活用が期待できる。

本章では響灘ビオトープに設置された 4K ネットワークカメラを対象に、チュウヒなど猛禽類を精度良くリアルタイムで検出できるよう、YOLO モデルの最適化を行ったうえで、鳥類検出動画の保存や飛跡追跡などの機能を開発し、その生活行動を観察するシステムの基盤構築について記す。また構築したシステムで鳥類の常時観察を行うことで、システムの課題抽出と改善を図りながら、運用方針を見出すことを目的とする。最終的な目標として、対象地区の最重要種であるチュウヒの生活行動を定量的に把握し、チュウヒ及びその採餌と営巣場所となる草地や湿地の生息環境の植生管理などの保全に貢献するシステムへの発展を目指す。

6.2 鳥類自動検出，追跡システム構築手法

6.2.1 撮影場所，撮影機材

動画の撮影場所は、福岡県北九州市若松区にある響灘ビオトープで、撮影機材はカメラ 4 台で全周撮影可能なパナソニック製ネットワークカメラ WV-X8570N（動画解像度 3840×2160）を用い、パナソニック製デジタルカメラ DC-FZ85（動画解像度 3840×2160）も援用した。動画のフレームレートはそれぞれ 15.00fps，29.97fps である。ネットワークカメラの設置地点と 4 台のカメラの撮影方向を図 6-1 に示す。映像は例えば 3 番カメラなら、「rtsp://ユーザ名:パスワード@ip アドレス/MediaInput/stream_1/ch_3」より取得できる。本研究ではチュウヒの飛来が多く確認されている 3 番，4 番カメラの映像を用いた。



図 6-1 撮影地点と撮影方向

6.2.2 深層学習法の入力パラメータ，神経回路網モデルの最適化

物体検出には Ultralytics 社の YOLO-v5¹²⁾を用い、学習モデルには同社より提供されている学習済みモデルを用いた。この学習には COCO データセット¹³⁾が用いられており、内訳は識別種数 80，画像数約 33 万，アノテーション数約 150 万となっている。本研究では YOLO-v5 に用意されている神経網モデルのうち、入力画像サイズ 640 ピクセル（以下 px と表記）のモデルを用いたが、パラメータ数の違いで 4 つ

のモデルが用意されている（表 6-1）。表 6-1 において、パラメータ数単位はメガを表しており、入力サイズの単位は px である。また mAP0.5 は物体検出の評価指標で、式(1)で与えられる IOU (Intersection Over Union) が 0.5 を超えたものを正解として求めた識別精度の全識別種平均である。YOLO-v5 に対し、響灘ビオトープでのリアルタイム検出に最適な入力パラメータと神経網モデルを検証した。計算機環境は、NVIDIA 製 GeForceGTX1070Ti と、GPU ライブラリ CUDA を用い、python でプログラムを作成した。

$$IOU = \frac{\text{正解領域と検出領域の重った面積}}{\text{正解領域と検出領域の和集合の面積}} \quad (1)$$

表 6-1 YOLOv5 各神経網モデルの性能

モデル名	パラメータ数(M)	入力サイズ(px)	mAP0.5(%)
YOLOv5s	7.3	640	36.7
YOLOv5m	21.4	640	44.5
YOLOv5l	47	640	48.2
YOLOv5x	87.7	640	50.4

6.2.3 検出物追跡システム、録画システムの開発

YOLO のフレーム単位の検出結果を用い、フレーム間で同一と思われる物体を、位置座標を基に推定し追跡するシステムを開発した。また、ネットワークカメラの映像時間は鳥類が撮影された時間と比べ長大であり解析を困難なものとしている。そこで YOLO を用いたリアルタイム検出をライブ映像に対して行い、検出が続く期間のみ映像を保存するシステムを開発し録画時間の縮小を図った。

6.2.4 開発したシステムの常時運用

構築した録画システムを NVIDIA 製 GeForceRTX3070 と CUDA を搭載した端末へ実装し、響灘ビオトープに設置した。システムは職員の出勤に合わせて手動で起動し（毎朝 8 時～8 時 30 分、但し休館日の火曜日は除く）、約 9 時間後に自動で終了する。保存された動画は週に 1 度回収し、解析することでシステムの課題を抽出した。常時運用は 2022 年 1 月 22 日から行った。

6.2.5 学習モデルの改善

YOLO-v5 の学習済みモデルの識別種には響灘ビオトープで検出される物体の多くは含まれていないため、常時運用で得られた物体検出画像を学習させることで学習モデルの改善を行った。学習はパラメータの初期値に、COCO を用いて学習された学習済みモデルの値を用いて行った。

6.3 システムの開発および常時運用

6.3.1 YOLO 法を用いた鳥類自動検出

(1) パラメータの最適化

撮影された映像と神経網モデル YOLO-v5l を用いて検出パラメータの 1 つである識別種判定の閾値 (conf_thres) の最適化を行った。このパラメータは検出時に物体を識別する際の確信度の閾値で 0~1 の値をとる。映像に対し物体検出を行うと、遠方の物体であるほど、小さい物体であるほど、映像に映る面積は小さくなり確信度は低下することを確認した。既定値の 0.25 では、このような面積が小さい鳥類を検出できたとしても、閾値を満たさず検出から除外されるケースがあったため、conf_thres 値は規定値より低く設定し 0.2, 0.1, 0.05 の 3 値で検出数の違いを調べた。図 6-2 は閾値ごとの検出数の比較を表と画像で行ったものである。図 6-2 より 0.05 が一番多く検出したが、特に 0.2 と 0.1 の間に検出数の開きが見られた。また、検出画像の比較より 0.2 では検出漏れが確認できる。この結果を受けて以降は conf_thres の値を 0.05 とした。また、80 ある識別種のうち鳥類を検出したものは bird, kite, horse の 3 種だった。horse は地上に降りた鳥類に反応していた。これより検出する識別種を 80 種からこの 3 種に限定し、検出物全てを鳥類とみなすようにした。

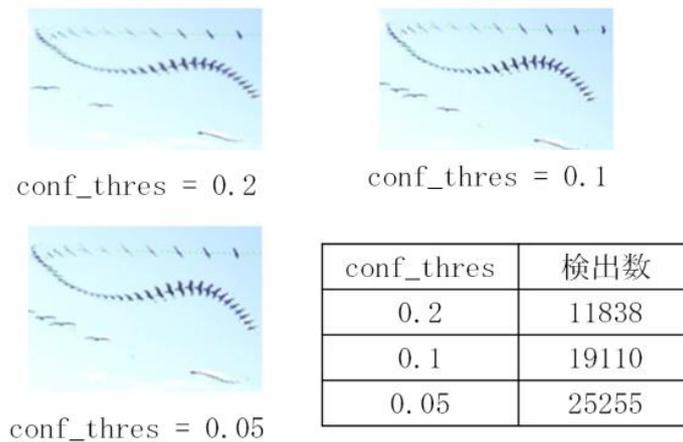


図 6-2 各 conf_thres 値の検出描画と検出数の比較

(2) 神経回路網モデルの最適化

響灘ビオトープで鳥類の検出を行う際に最適な学習モデルについて、各神経回路網モデルの検出時間と検出数を調べた結果を表 6-2 に示す。表 6-1 と表 6-2 より、パラメータ数の増加によって検出時間は増加しているが、検出数の大きな差異はなかった。リアルタイム検出の可否については、表 6-5 より 15fps のフレームレートでは全てのモデルでリアルタイム検出が行える。検出数は高性能モデル v5x よりも、低性能モデル v5m, v5l の方が多いが、これは検出対象の姿形の不明瞭さが一因と考えられる。以上の検出数と検出時間から、用いる神経網モデルに YOLO-v5m を選定した。また、これらの最適化により、4K カメラ映像に対し検出物の最小面積はピクセル数にして 156px であった。

表 6-2 YOLO-v5 各モデルの検出結果

モデル名	平均検出時間 (s)	検出数
YOLO-v5s	0.008	431
YOLO-v5m	0.015	532
YOLO-v5l	0.023	524
YOLO-v5x	0.04	487

6.3.2 追跡システム、録画システムの開発

(1) 固定検出物の除去

追跡システム構築に先立ち誤検出物の削除を行った。図 6-3 左図は奥行きを時間軸とした、検出物の検出位置のプロット図を示すが、奥へ伸びる 2 本の帯状の検出物が確認できる。これは特定地点で物体を検出し続けていることを示すが、検出画像から風車の羽を誤検出していることが分かった。この誤検出は `conf_thres` を下げた影響と考えられ、検出時ではなく検出後に除去する必要がある。移動する鳥類に対し、風車は固定されているので、固定検出物を削除するプログラムを独自に作成した。その機能は、映像を縦横 5px の矩形で分割し、矩形内の検出数が 5 回を超えた場合、その領域の検出物を固定検出物として削除する。本プログラムを適用した結果を図 6-3 右図に示すが、同左図との比較により固定検出物の多くを除去できていることが分かる。

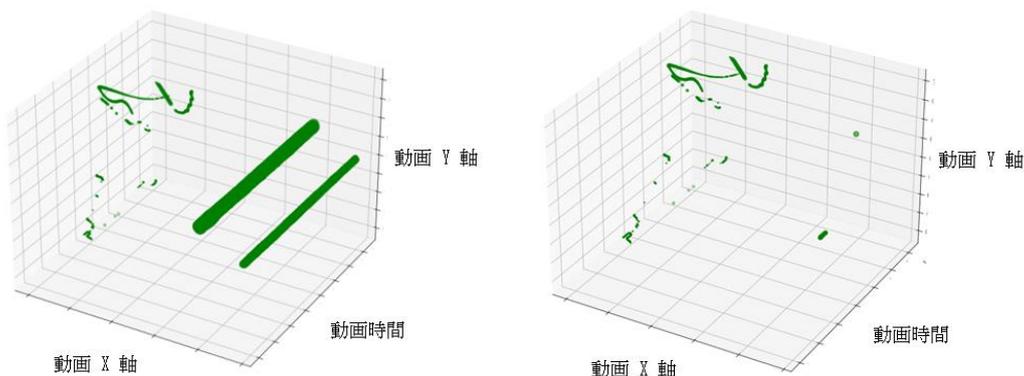


図 6-3 検出物の検出位置プロット図 (左) と固定検出物削除後のプロット図 (右)

(2) 検出物追跡システムの構築

検出物を追跡する手法は、移動速度を保持した検出物のグループ化により行った。対象検出物に対しその大きさの 5 倍の探索領域を設定し、探索領域内にある前フレーム検出物の重心位置の中から、対象検出物の重心位置が最小となるものを同一物と見なしグループ化する。この時、前フレーム検出物はグループ化されており、その位置はグループが持つ速度を足した値を用いる。速度の初期値は 0 とし、グループに検出物が追加される毎に前フレームからの移動量に更新する。これにより検出物の追跡を行うが、追跡できなかった場合は探索領域を 5 倍から 10 倍に拡張して同様の手法で探索し、50 フレーム追跡できない場合は追跡を中断するようにした。図 6-4 にこのプログラムを用いて検出物の追跡を行い、その飛跡を描画した結果を示す。中央の飛跡の途切れを除き、正確に追跡できていることが確認できる。また別の映像では 2 羽が交差するように飛行しており、



図 6-4 追跡プログラムによる飛跡描画



図 6-5 交差する飛跡の描画結果

交差点で間違わずに追跡できていることが確認できる(図 6-5)。これは速度の考慮により追跡点を特定できた結果と言える。これにより YOLO-v5 を用いて映像から鳥類を検出し、それを自動で追跡する飛跡解析が可能になった。

(3) 検出時録画システムの構築

YOLO-v5 の動画リアルタイム検出機能を利用して、ネットワークカメラのライブ映像から鳥類が検出された場合に映像を保存できるシステムを構築した。カメラの fps は 5.0~15.0 に設定できるが、低レートでは高速で飛行する鳥類を捕捉できないため、最大値の 15.0 に設定した。これにより得られた映像に対し最適化した YOLO-v5 を用いて物体検出を行い、検出があった場合は未検出が 2 分間続くまで映像を保存し、未検出が続く時間の映像は保存しないシステムを構築した。人を対象とした本システムの検証では 480 分の撮影時間に対し録画された映像の合計時間は 294 分で、約 60%の圧縮を確認した。

6.3.3 システムの常時運用

響灘ビオトープにて構築したシステムの継続運用を、2022 年 1 月 22 日からの試験運用を経て 2022 年 2 月 1 日より開始した。2022 年 8 月 31 日までの期間で録画が確認された日数は 139 日だったが、設定した 9 時間までシステムが稼働していない日も多く確認された。これはシステム全体の処理時間が 1 フレームの時間を超えたことや、録画時間の長期化によるメモリの圧迫などが原因と考えられる。図 6-6 に運用期間中に検出された鳥類の一例を示す。図 6-6 右下の遠方で検出された鳥類は、旋回する様子が確認できたので猛禽類と推定した。図 6-6 より、姿形の明瞭さを問わない検出、個体と群体、その両方の検出を確認できた。さらに検出物は、図 6-6 上部 2 枚のように姿形が明瞭である場合はそれを正確に捉え、図 7 右下のように造詣が不明瞭な場合でも



ツバメ トビ
カルガモの群れ 猛禽類(種の識別不可)

図 6-6 システム運用にて検出した鳥類の一例

その位置を正確に捉えていることが確認できる。運用期間における検出物の最小面積はピクセル数にして 320px であった。

また季節による検出の違いについて解析を行った。冬季，春季，夏期（2月，4月，8月）において，8時～17時の間検出が行われた日のうち連続した3日間を選び，検出時刻別と検出鳥類の大きさ別で整理した。大きさは大型，中型，小型の3種で，これは検出物自体の大きさではなく，検出位置，飛跡などから推定した種の大きさである。大型にはトビをはじめとした猛禽類を，中型にはカラス，ハト，カモ類を，小型にはスズメ，ツバメ，オオヨシキリなどを含めた。これらの判断は飛行位置，飛び方，飛行速度などから人が判定している。解析した期間はそれぞれ2月23日～2月25日，4月9日～4月11日，8月19日～8月21日で，大きさ別に整理した結果を表6-3に示す。また図6-7に時間別の検出数をグラフで示す。表6-3より8月の検出数が1番多く2月と比べ約3倍，4月と比べ約1.8倍となった。大きさ別に見ると8月の小型鳥類の検出が顕著に増加している。これらの多くは図6-6左上に類似した姿形と飛跡であることを確認しており，8月にはツバメの活動が活発になっていると推察できる。中型，大型については8月の検出数は他の月と比べ減少しているが，小型のように顕著ではなくより長い期間での解析が必要である。時刻別に検出数を比較すると図8より2月は10時台，4月は13時台，8月は12時台に検出のピークを迎えている。いずれの時間帯も日中の活動期にあたることから，多くの鳥類がビオトープで何らかの活動をしていると推察される。

さらに2月23日について1日の飛跡分布を解析した。その結果を図6-8に示すが，多くの鳥類が移動飛行時に検出されており，採餌行動などは確認できなかった。画像下部の飛跡が画像上部に比べジグザグになっているが，近距離で検出されたため羽ばたきの影響から検出座標が上下したものと考えられる。

表 6-3 大きさ別検出数

	大型	中型	小型	合計
2月	48	42	8	98
4月	49	58	59	166
8月	32	22	250	304

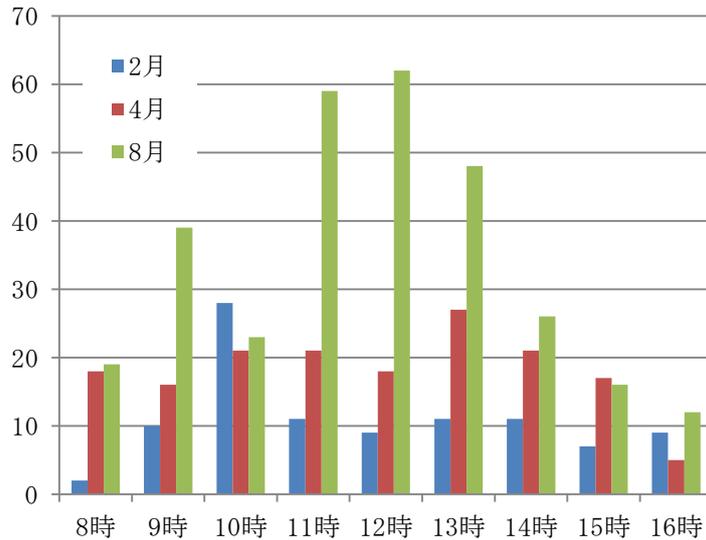


図 6-7 時間別検出数の比較

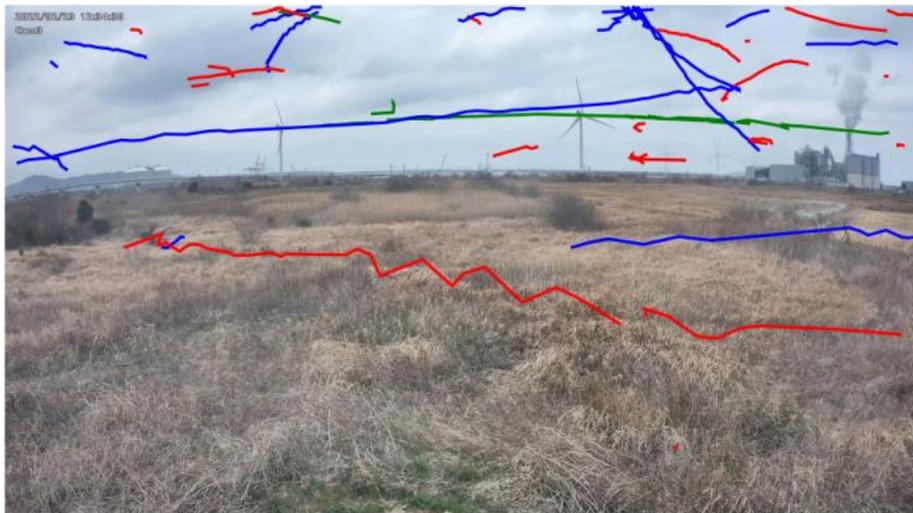


図 6-8 2月23日の飛跡プロット

システム運用の課題としては稼働時間安定化の他にも誤検出が挙げられる(図 6-9)。誤検出された物体は風車、煙、雲、植物、影などで、春季以降は昆虫の誤検出も確認された。昆虫以外は動きの少ない固定検出物であるため、作成したプログラムを用いて除去できるが、昆虫の誤検出は手作業で除去する必要がある。また録画システムによる録画時間の削減については、その日の録画開始時刻から録画終了時刻までの時間とその日の総録画時間とで求めた平均圧縮率が 47.7%であった。誤検出の低減により圧縮



図 6-9 常時運用において確認された誤検出物

率はさらに向上するものと考えられる。

6.3.4 学習モデルの改善

誤検出の課題解決には学習モデルの改善が必要と考えられるため、4つのデータセットを作成し比較を行った。作成したデータセットを、COCO データセット (A)、A から airplane, bird, horse, kite を抽出したもの (B)、B に常時運用で誤検出が目立った風車、植物、雲・煙を追加したもの (C)、C の bird にビオトープで検出された鳥類画像を追加したもの (D) とし、学習結果の比較を表 6-4 に、検出結果の比較を表 6-5 にそれぞれ示す。B の抽出種は 6.3.1 において鳥類と識別した 3 種に、遠方を飛行した際の誤検出の可能性を考慮し airplane を加えた。検出には 400 秒の動画を用いた。表 6-4 において train, val はそれぞれ学習と検証に用いたアノテーション数で、表 6-5 において検出数は鳥 (bird) と凧 (kite) の検出数を合計したもので、各種の検出数は人が判定している。表 6-4 より識別種が 80 ある A に対し種数の少ない他のモデルは mAP0.5 が約 10% 向上している。表 6-5 より種数を限定した B は A に対し風車の誤検出が減少し、鳥類の検出が増加した。C では誤検出が劇的に改善した一方で、鳥類の検出数も A と B の検出数の 30%~46% に止まり、検出漏れが発生している。D では風車と煙・雲の誤検出は低いままだが、植物の誤検出が劇的に増加した。これは草陰を草地に降り立った鳥類と判断した結果と考えられる。鳥類の検出に関しては B と比較して約 2.3 倍に増加し、遠方や羽を畳んだ鳥類の検出性能が向上している。

表 6-4 作成したデータセットの学習結果

データ名	種数	train	val	平均 mAP
A	80	849942	36336	64.1
B	4	31040	1169	76.5
C	7	32311	1488	74.4
D	7	35003	2162	76.5

表 6-5 作成したデータセットの検出結果

データ名	検出数	風車	植物	煙・雲	鳥類
A	1328	1039	26	15	231
B	1336	776	165	29	353
C	116	4	1	0	107
D	2468	5	1613	3	802

6.4 議論

本研究で得られた最小検出面積は 156px で 4K 画像面積 8294400px に対して 0.002% 以下であり、物体の矩形形状は正確に抽出できた。このことより、大きさが分かっている鳥類に対しては、検出した大きさより距離を見積もることができる。翼開長約 1.5m のトビなら、距離が分かっている風車とのサイズ比から検出面積 418px の場合、距離は約 280m と推定できる。156px なら 458m となる。4K 画像は YOLO へ入力する際に画像サイズは 1/3 (横幅 3840 から 1280) に縮小されるので、4K 画像から 1280 サイズの部分画像を切り出せば画像の縮小が発生せず、3 倍の距離 (トビで約 1.3km) まで離れても検出できることになる。これはビオトープでカメラの増設・配置計画に活用できる。

学習モデルの改善において、ビオトープで撮影した画像を加えることで鳥類の検出性能は大幅に向上し、風車や煙、雲といった誤検出を大幅に低減することが確認できた。COCO データセットのアノテーション数に対し本研究で追加したアノテーション数は少ない。YOLO-v5 においてよい結果を得るために必要なアノテーション数は識別種 1 種につき 1 万以上と言われているが¹⁴⁾、汎用性を考慮しない場合に必要なアノテーション数は低下すると考えられる。システム運用において鳥類検出での成果は得られているので、今後も学習モデルの改善を反映させながら運用を続け、アノテーションを追加していく。さらに識別種数、学習に用いる COCO データセットの比率、conf_thres 値などを調整し響灘ビオトープに最適な学習モデルの構築を図る。

本研究では飛跡を平面に描画することができたが、1つのカメラ映像から得られる平面情報では鳥類の生活行動把握には不十分である。複数のカメラを用いたステレオ撮影は飛跡の 3次元解析に応用されており¹⁵⁾、響灘ビオトープにおいてもネットワークカメラの複数箇所への設置は計画されている。今後は本研究で得られた結果や本研究では用いなかった 2つの方角の映像も活用し、より詳細な鳥類の生活行動把握のための手法について検討を進める。また本研究では扱えなかった検出鳥類の種の識別について、検出に用いた YOLO 法は映像内面積が小さい鳥類を検出するため確信度閾値を下げており、鳥類種識別への利用は確信度の観点から困難である。そこで鳥類種の識別には深層学習法による画像単位の識別と連結させる方法が有効と考えられる。例えば第 4 章と第 5 章で一定の成果を得た神経回路網モデル Xception¹⁶⁾を用いて、Xception の入力サイズ 299×299 近辺を下限とし、検出物を学習画像として学習させることで、下限サイズ以上の検出面積を持つ鳥類の種の識別が可能と考えられる。

6.5 本章のまとめ

本研究では、深層学習法 YOLO を用いて 4K 動画のリアルタイムでの鳥類の自動検出、および検出鳥類の自動追跡システムを構築し、常時運用を経てシステムの課題抽出と改善を図りながら運用方針を見出した。本研究で得られた知見、指針を以下にまとめる。

①4K 動画でも YOLO-v5m モデルでリアルタイム検出が可能で、面積比 0.002% 以下の小さな物体まで精度よく

検出できる。

②YOLO-v5 提供の学習済みモデルは識別種が 80 種類と少なく誤認識は多いが、物体の矩形形状は正確に抽出できるため、アノテーションへ流用することで容易に独自学習用の画像データが作成できる。

③YOLO-v5 の学習モデルに、1 割程度の独自画像を加えるだけでも認識精度は大きく改善する。利用目的に最適な識別種の選定、および学習用画像データセットの整備が重要である。

参考文献

- 1) 響灘ビオトープ, http://www.hibikinadabiotope.com/?page_id=131, 参照 2023-01-06
- 2) 環境省 (2020) : 環境省レッドリスト 2020, <https://www.env.go.jp/press/107905.html>, 参照 2023-01-06
- 3) 環境省 (2015) : 平成 26 年度チュウヒ保護方策検討委託業務報告書
- 4) 福岡県 (2011) : 福岡県レッドデータブック 2011【改訂版】, <http://www.fihes.pref.fukuoka.jp/~kankyouseibutsu/rdb/rdb.html>, 参照 2023-01-06
- 5) 北九州市 : 環境影響評価の進捗状況, <https://www.city.kitakyushu.lg.jp/kankyou/00600320.html>, 参照 2023-01-06
- 6) 濱尾章二 (2011) : 鳥類の多様性を把握するための調査方法の検討:ラインセンサス法と捕獲法の比較, 自然教育園報告, No.42, pp.1-12
- 7) 山川将徑・北村俊平 (2016) : 石川県林業試験場の森における鳥類記録手法 (音声録音とスポットセンサス) の比較, 石川県立自然史資料館研究報告, No.6, pp.15-24
- 8) 田悟和巳・高橋・明寛・萩原陽二郎・益子理・横山陽子・近藤弘章・有山義昭・樋口広芳 (2020) : レーダーを用いた夜間の渡り鳥の飛跡数, 飛翔高度, 渡り経路の追跡, 日本鳥学会誌, Vol.69, No.1, pp.41-61
- 9) Joseph, R., Santosh, D., Ross, G. and Ali, F. (2016) : You Only Look Once: Unified, Real-Time Object Detection, Computer Vision and Pattern Recognition Conference 2016, pp.779-788
- 10) 北村聡史・中山紘喜・野上敦嗣 (2019) : ステレオ法による 3 次元動体追跡システムの開発, 環境システム研究論文発表会講演集, Vol.47, No.1, pp.49-54
- 11) Grabner, H., Grabner, M., and Bischof, H. (2006) : Real-Time Tracking via On-Line Boosting, British Machine Vision Conference
- 12) YOLOv5, <https://github.com/ultralytics/yolov5>, 参照 2023-01-06
- 13) COCOdataset, <https://cocodataset.org/#home>, 参照 2023-01-06
- 14) YOLOv5 Documentation, <https://docs.ultralytics.com/tutorials/raining-tips-best-results/>, 参照 2023-01-06
- 15) 村井祐一・田坂裕司・武田靖 (2009) : 応用ステレオ PTV による大気中の飛翔体の 3 次元計測, 可視化情報学会誌, Vol.29, No.1, pp.371-372
- 16) Francois, C. (2017) : Xception: Deep Learning with Depthwise Separable Convolutions, Computer Vision and Pattern Recognition Conference, 2017, pp.1251-1258

第7章 議論

第4章～第6章までで深層学習法を用いた生物調査手法を構築、運用しその適用性を評価した。画像識別と物体検出においては高精度なシステムを開発できたが、分布図の作成即ち領域分割においては、区分精度、識別精度ともに課題の残る結果となった。区分精度については木本種、草本種双方において複数種の枝葉が絡み合い群生している場合が多く、本研究で用いた MeanShift 法では必ずしも種ごとの明確な領域分割ができるとは限らない。このことから区分精度の改善には、分割サイズを縮小し複数種が群生している場合でも、なるべく種ごとの区分ができるようにデータ増強を行う方針と、複数種の群生を新たな識別種として登録しデータ増強を行う方針、これら2つの方針が考えられる。しかし前者の方針は、対象物との距離が離れる広域撮影において、分割サイズが更に縮小され区分精度改善の障害になる可能性が高い。後者の方針においても複数の識別種に該当する種が発生し、識別精度への影響が懸念される。識別精度については第4章、第5章で構築した学習データの多くは、分布図作成に用いた画像とは異なる距離、角度から撮影された画像であることが、正しく識別できない一因と考えられる。今後、学習データの増強による精度の変化を検証する必要性は高い他、画像単位の識別でなく画素単位で識別を行う領域検出（セマンティックセグメンテーション）の適用について検証を行う。

第6章において、公開されている汎用的な学習データへ、識別や検出を行う映像に近い環境で得られたデータを加えることで精度の改善が確認された。これより学習データの増強にはオリジナルデータを増強する仕組みを整えることが有用と言える。撮影方針として、対象を大きく捉え被写界深度を狭めることが有効との知見を得たが、これは CNN が一般的に位置不変性を持ち、対象の位置、スケール、回転の変化に強いことにも合致している。しかし植物分布図作成において、作成した画像データベースには改善の余地があり、適切なスケールと画像数に対するその割合を検証する必要がある。また種によっては、発見数の違いから学習画像数を十分に確保できないケースがある。第4章、第5章では転移学習により、50枚程度の画像数でも高い確度が得られる種もあったが、画像数の少ない種に確度低下の傾向がみられ、画像数の変化が別種の確度にも影響していることを第5章で確認した。第4章では種ごとに確度低下の要因解析を行ったが、画像数や種の全体に対する画像割合については検証の余地がある。

本研究で構築した植物識別システムは、調査によって得られた画像を用いた深層学習により学習モデルを作成し、湿地、台地、河川敷、里山において一定の成果を得た。他地区でも同様に調査より集めた画像を基に深層学習を行うことで、精度の高い学習モデルの構築が可能と考えられる。ただし人による立ち入り調査が行えない場合、第4章で実施した人手による背景のぼかしが行えないので、イネ科などの細かい植物種の識別が課題となる。また鳥類検出システムについては、カラスやハト以上の大きさであれば検出頻度のよい結果が得られると考えられるが、小型鳥類については検出面積が小さくなりがちなので、検出頻度や検出距離が課題となる。また本研究では見晴らしの良い湿地にカメラを設置したため、視界を確保できたが、森林などで実施する場合は視界の確保を考慮する必要がある。さらに2つのシステムから得られたデータを連携させることで、保護種の生活環境を整える保全策への適用も考えられる。響灘ビオトープではチュウヒの保全を行っているが、植物分布図よりチュウヒの採餌や繁殖に適した植生を整理し、その経過を鳥類検出によりモニタリングするなどの応用が考えられる。

鳥類においては第6章で構築したシステムの運用を続けることで、オリジナルデータの増強は可能である。植物においては第5章で活用したドローン撮影が、短時間で広域の撮影が可能であり有用と考えられる。これら収集したオリジナルデータは、深層学習法による解析結果と併せて、環境学習への活用

が期待できる。響灘ビオトープおよびその周辺地区は、エコタウン事業や低炭素社会に向けた取り組みをはじめ、持続可能な社会づくりに求められる 3 要素（低炭素、資源循環、自然共生）が全て揃ったアジアのグリーンショールームを目指した先進的な取り組みが集積する場所でもあり、環境教育や環境情報の発信を積極的に行っている。高精度な植物分布図は視覚資料として有用であり、常時観測により蓄積されたデータは説得力を持った資料として有用である。このような取り組みは、保全緑地の維持管理において課題となっている、重要種保全の必要性の認知に対し課題解決の一助となり、緑地保全の維持管理に寄与できる。また精度の高い樹木分布図を作成できれば、対象地区の CO2 吸収量など森林機能の推定へ応用でき、都市計画、保全計画、環境学習などへの活用が期待できる。このような側面もあり、今後精度改善へ向けた継続的なオリジナルデータの増強と、深層学習法を用いた種々の解析及び検証の継続には意義がある。

第8章 結論

8.1 各章のまとめ

本研究は、北九州市若松区に位置する響灘ビオトープを中心に、保全緑地の生物種に対し深層学習法を用いた生物調査手法を構築、運用しその適用性を評価したものである。以下にその成果をまとめる。

第1章では、保全緑地の重要性とその維持管理が抱える課題について述べるとともに、深層学習法や小型UAVといったIT技術活用の有用性について述べ、本研究の目的を明確にした。

第2章では、生物保全分野における深層学習法の応用研究に関する既往研究のレビューについて述べ、本研究の位置づけを定めた。

第3章では、本研究手法において主要な技術である畳み込みニューラルネットワークとYOLOによる物体検出の仕組みについて述べ、その動作原理についてまとめた。

第4章では、響灘ビオトープ、ひびきの北公園、撥川において植物調査を行い、多植物の識別における精度改善と汎化性能の向上を目的とし全10回の植物調査を行い、Xceptionによる植物識別の学習モデルを構築した。結果として121種の植物画像のデータセットを作成し、98%（上位3位では99.5%以上）の高い確度で分類することが出来た。学習後試験において植物種によって確度が大きく低下した植物についても要因を解明することができた。また要因解析に基づき、新たな植物調査と学習を行い、汎化性能を高め、新たに確度が低下した植物は要因解析を行うことが出来た。

第5章では、木本種を対象とし、①深層学習法を用いた樹冠識別から植生分布を求める手法の構築、②里山樹木に対して、高い確度を得る画像群の撮影条件の検討および本システムの里山管理への適用について、樹木分布図の作成を目的とし里山樹木の調査を行った。その結果デジタルカメラとドローンを用いて、多様な角度で撮影したオリジナル画像から樹木8種を高精度に識別する学習モデルを作成できた。8種のうち樹冠を撮影できた7種について樹冠識別を行うと、Top3の確度は約50%だったが、アスペクト比の統一や類似した部分画像の排除に加え、試験画像と同程度の距離と角度で撮影した画像を含めた学習を行うことで、7種のうち6種の確度が約90%に向上した。構築した学習モデルを用いて樹木分布図を作成し、草本種においても同様に分布図を作成したが、分割精度や識別精度において一定の成果を得ると共に課題と解決策について検討することができた。

第6章では、鳥類調査に対する深層学習法の応用において、ビデオ映像からの鳥類自動検出および自動追跡について検証した。YOLO法を用いて4K動画のリアルタイムでの鳥類の自動検出および検出鳥類の自動追跡システムを構築し、常時運用を経て鳥類の生態解析とシステムの課題を明らかにした。検出パラメータの最適化により4K動画に対し最小で0.002%以下の鳥類を検出できた。また構築したシステムの常時運用で得た画像から学習用データを作成し独自に学習したモデルを用いることで鳥類の検出精度の向上と誤検出の大幅な低減を達成できた。

第7章では、第4章～第6章で得られた結果を基に研究課題と今後の展望についてまとめた。

8.2 本論文の結論

本研究の結論を以下にまとめる。生物の調査や保全への深層学習法による教師あり学習の適用性について、植物と鳥類を対象に物体分類や物体検出を行うシステムを構築し、評価を行った。オリジナル画像とその部分画像から作成した画像データベースと転移学習により、植物識別において高い確度を達成し、公開されている画像データベースにオリジナル画像を加えた学習により、鳥類検出において高い検

出頻度と正確な検出を達成できた。この成果より、撮影手法、画像データベースの構成、深層学習法の性能限界について知見を得ることができ、深層学習法による教師あり学習は生物の調査や保全に対し、十分な適用性を持っていると言える。一方で、植物分布図や飛跡解析といった解析については課題があり、撮影手法や画像データベースの構成について更なる検証が必要である他、領域検出、飛跡の3次元解析といった、本研究では扱うことができなかった手法によるアプローチも必要である。構築したシステムは複数の環境で実施しており、他地区での運用も十分に考えられる他、得られたデータから前述したような解析の高精度化が実現すれば、環境保全や環境教育、都市計画などへの応用も考えられる。

謝辞

本論文は、北九州市立大学大学院国際環境工学研究科環境システム学専攻の博士後期課程に在籍中に実施した研究を取りまとめたものです。取りまとめにあたって、たくさんのご指導、ご協力を頂きました。ここに謝辞として記させていただきます。

主査を引き受けていただきました松本 亨教授には、1年という短い期間の中で、査読付き学術論文の作成から博士論文の取りまとめに渡り、多くのご助言とご指導をいただきました。最後までご心配をおかけしましたが、無事に博士論文を取りまとめることができました。深く感謝申し上げます。

大変お忙しい中に副査を引き受けていただきました福田 展淳教授、大矢 仁史教授、加藤 尊秋教授には、本論文の審査を通して貴重なご助言をいただき、博士論文が論ずべき点など多くを勉強させていただきました。深く感謝申し上げます。

白石 瑠菜さんをはじめ、野上研究室に在籍していた皆様には、現地調査、学術論文の作成など研究面で多大な協力をしていただきました。心より感謝を申し上げます。

安枝 裕司様をはじめ、響灘ビオトープの皆様、久富 学様をはじめ、株式会社エコプラン研究所の皆様には現地調査、解析作業において多大なご協力、ご助言をいただきました。心より感謝を申し上げます。これからもお世話になると思いますがよろしく申し上げます。

2020年4月より、ヒューマンアカデミー株式会社、福岡校へ契約講師として業務に携わらせていただきました。研究との兼務の中、安本 良枝様をはじめ研究活動の重要性をご理解いただき、ご協力いただきましたこと厚く御礼申し上げます。

社会人大学院生の道を6年にわたり、生活面、経済面、精神面から支え、応援してくれた家族に心から感謝しています。ありがとうございました。

最後に、2022年3月をもって北九州市立大学を退職なされた野上 敦嗣様、社会人大学院生として研究の機会を与えてくださり、研究テーマの設定から学術論文、博士論文の構成、研究者としての心構えなど丁寧にご指導いただきました。要領の悪い私は先生に多くのご心配とご迷惑をおかけしましたが、最後まで熱心にご指導ご鞭撻いただきましたこと深く感謝いたします。本当にありがとうございました。

2023年3月13日

中山 紘喜